

DOCUMENT RESUME

ED 259 909

SE 045 880

AUTHOR Fennell, Francis, Ed.
 TITLE The Research Council for Diagnostic and Prescriptive Mathematics 1983 Research Monograph. Selected Papers from the Sixth, Eighth and Ninth Annual Conferences on Diagnostic and Prescriptive Mathematics (Tampa 1979, Hershey 1981, and Buffalo 1982).
 INSTITUTION Research Council for Diagnostic and Prescriptive Mathematics, Kent, OH.
 REPORT NO ISBN-0-940466-08-2
 PUB DATE 83
 NOTE 141p.
 AVAILABLE FROM RCDPM Headquarters, 441 Beryl Dr., Kent, OH 44240.
 PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143) -- Collected Works - General (020)

EDRS PRICE MF01/PC06 Plus Postage.
 DESCRIPTORS Cerebral Dominance; *Diagnostic Teaching; Diagnostic Tests; Educational Research; Elementary School Mathematics; Elementary Secondary Education; *Mathematics Achievement; *Mathematics Education; *Mathematics Instruction; *Problem Solving; Secondary School Mathematics; *Teaching Methods; Testing

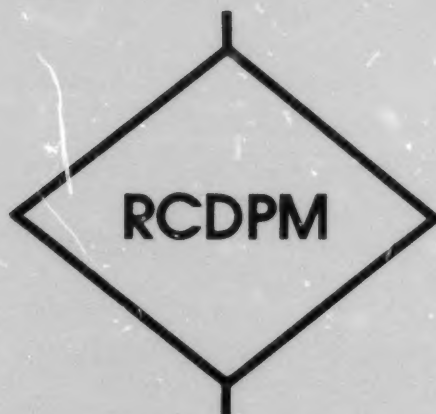
IDENTIFIERS *Mathematics Education Research

ABSTRACT

Papers from three annual conferences of the Research Council for Diagnostic and Prescriptive Mathematics are presented. They are: "The Mathematics Clinic Model: Parents as Clinicians and their Effect on Achievement, Attitude, and Self-Concept" (Lee Wise and Janet Scheer); "Heterogeneous versus Homogeneous Item Arrangements in Mathematics Testing" (John Engelhardt, Linda Oldaker, and Leonard Cahen); "Testing Mathematical Problem Solving" (Thomas Romberg, Kevin Collis, and Mirad Jurdak); "A Model and Sequence for Diagnostic Instruction in Problem Solving" (Francis Fennell); "The Validation of a Criterion-Referenced Test of Rational Number Concepts and Skills" (Barbara Sadowski); "An Exploration: White and Clark's Test of Inclusion Applied to Ex Post Facto Learning Data" (Bob Underhill); "Brain Hemispheric Considerations for a Developmental Approach to Diagnosing Mathematics Concepts for Children" (Janice Flake); "Doing it Their Way: Some Child-Methods in Mathematics" (Lesley Booth and Kathleen Hart); "Diagnostic and Prescriptive Instruction Applied to Teaching Mathematically Talented Students" (Kevin Bartkovich); "Linguistic Aspects of More/Less and Dimensional Adjectives Relevant to Diagnosis and Instruction" (Carol Novillis Larson); "Improving Computation--A look at Estimation and Approximation as Significant Elements in Achieving Understanding and Correctness" (Douglas Crawford); and "Strategies for Diagnosing and Enhancing Learning of Place-Value Concepts" (Bob Underhill). (JN)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED259909



U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

✓ This document has been reproduced as
received from the person or organization
originating it.

□ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

1983 RESEARCH MONOGRAPH

Selected Papers From the
Sixth, Eighth, and Ninth National Conferences on
Diagnostic and Prescriptive Mathematics

Managing Editor
Francis (Skip) Fennell
Western Maryland College

Copy Editors:
Jane M. Hill
National Council of Teachers of Mathematics

Grace Burton
University of North Carolina - Wilmington

Jan Scheer
Southern Illinois University

RESEARCH COUNCIL
FOR
DIAGNOSTIC AND PRESCRIPTIVE MATHEMATICS

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Leonard M. Kennedy

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

SE 045880

COPYRIGHT, 1983, BY THE RESEARCH COUNCIL FOR DIAGNOSTIC AND
PRESCRIPTIVE MATHEMATICS.
FIRST PRINTING, MARCH, 1983, AT THE STATE UNIVERSITY OF NEW YORK AT BUFFALO.

PRINTED IN THE UNITED STATES OF AMERICA

ISBN 0-940466-08-2

LC 83-060734

ALL RIGHTS RESERVED. NO PART OF THIS MONOGRAPH MAY BE REPRODUCED OR TRANSMITTED IN ANY FORM OR BY ANY MEANS, ELECTRONIC OR MECHANICAL, INCLUDING PHOTOCOPY, RECORDING, OR ANY INFORMATION STORAGE AND RETRIEVAL SYSTEM, WITHOUT PERMISSION IN WRITING FROM THE PUBLICATION CHAIRPERSON.

PUBLICATIONS OF RCDPM ARE PRODUCED UNDER THE AUSPICES OF THE PUBLICATIONS COMMITTEE OF THE ORGANIZATION. THE 1982-83 COMMITTEE MEMBERS WERE: CAROL NOVILLIS LARSON, CHAIRPERSON; JAMES BRUNI; DOUGLAS BRUMBAUGH; HELEN CHEEK; FRANCIS (SKIP) FENNELL; AND JANE M. HILL.

THE MATERIALS PRINTED IN THIS MONOGRAPH ARE THE EXPRESSION OF THE WRITERS AND NOT NECESSARILY A STATEMENT OF POLICY OF THE RESEARCH COUNCIL FOR DIAGNOSTIC AND PRESCRIPTIVE MATHEMATICS.

ADDITIONAL COPIES OF THIS OR ANY PUBLICATION OF RCDPM MAY BE PURCHASED FROM RCDPM HEADQUARTERS, 441 BERYL DRIVE, KENT, OHIO 44240.

THE RESEARCH COUNCIL FOR
DIAGNOSTIC AND PRESCRIPTIVE MATHEMATICS
1983 RESEARCH MONOGRAPH

Selected Papers from the
Sixth, Eighth and Ninth Annual Conferences
Tampa 1979 Hershey 1981 Buffalo 1982

Managing Editor

Francis (Skip) Fennell
Western Maryland College

Production Editor

Leroy G. Callahan
S.U.N.Y. at Buffalo

Printing and Distribution

Leroy G. Callahan
S.U.N.Y. at Buffalo

James Heddens
Kent State University

Copy Editors

Jane M. Hill
NCTM

Grace Burton
University North Carolina-Wilmington

Jan Scheer
Southern Illinois University

Vice President for Dissemination

Linda Oldaker Gemill
Arizona State University

Preface

As always the major purpose of preparing this 1983 research monograph is to make available to RCDPM members and others the significant thoughts, ideas, premises, research hypotheses, and conclusions of a recent RCDPM conference.

However, this particular research monograph is different for varied reasons. First, it presents papers from the Tampa 1979, Hershey 1981, and Buffalo 1982 RCDPM Conferences, so this monograph truly samples the myriad of diagnostic/prescriptive interests presented at the 6th, 8th, and 9th conferences. A second difference of this monograph is the fact that each of the twenty-six manuscripts considered for publication was reviewed by a minimum of three of the forty RCDPM referees used to guarantee quality, professional writing within our council. As managing editor of this publication, I can't possibly say enough about the careful reading, editing, attention to APA style guidelines, and genuine concern of our referees. A personal acknowledgment follows this preface to formally thank the 1983 monograph referees.

Perhaps it's time for all council members to visualize the shaping of the research publication process as it now exists within RCDPM. Under the direction of the Chairperson of the Publications Committee within RCDPM, Carol Novillis Larson, papers were submitted and forwarded to me. Each paper was then reviewed and a decision made as to its publication in this RCDPM monograph, the RCDPM newsletter, revision, or rejection. Accepted papers were rewritten and copy-edited. Papers to be revised all went through a second refereeing. All papers were copy-edited by either Jane M. Hill, of NCTM, Grace M. Burton, of the University of North Carolina-Wilmington, or Janet Scheer of Southern Illinois University. Let there be no doubt that Hill, Burton, and Scheer can cut any author's verbosity to precision! Once papers were copy-edited, authors had to complete a final draft, which came under the scrutiny of this writer's sometimes fuzzy eyes. Finally, papers were sent to Leroy Callahan of the SUNY at Buffalo for production. Roy Callahan's title of Production Manager is somewhat misleading. What our manager did was make sure the entire monograph was re-typed, edited, now a fifth time for some papers, and ordered appropriately. Harry Truman's classic statement, "the buck stops here" fits Roy Callahan's integral role in this publication.

Lest I forget, and while recognizing that it may be overstepping my editorial license, Peg Callahan deserves more than just an acknowledgment for typing this monograph. Papers with topics, figures, tables, and issues as different as test construction, gifted children, parent involvement, and learning theory fill the following pages. Peg, thanks so much.

While I'm extending kudos, I would also like to acknowledge Linda Oldaker Gemmill's leadership as Vice President for Dissemination. Linda truly aided in keeping the authors of Tampa and Hershey papers hopeful about the eventual publication of their work.

1983 monograph papers from the 1979 Tampa RCDPM Conference include Janice Flake's review of research regarding child development and Carol Novillis Larson's paper which focuses on relational terms and their use.

Papers from the 1981 Hershey RCDPM Conference include those by Engelhardt, Cahen, and Oldaker-Gemmill, and Sadowski relating to testing and diagnosis. Crawford's paper on approximation and estimation and Fenrell's on problem solving suggest that diagnostic/prescriptive techniques readily apply to these rather topical thrusts in mathematics education. Underhill's continuing work regarding the validation of learning hierarchies is continued with his paper regarding White and Clark's Test of Inclusion.

1982 Buffalo Conference papers include Kevin Bartkovich's which presents a model for teaching gifted and talented students. Scheer and Wise collaborated on a research paper related to parents and diagnosis. Romberg's paper based on item construction and Underhill's Soviet teaching experiment are research attempts of interest to all of us. Lastly, Leslie Booth and Kath Hart provide some interesting research regarding child methods for solving problems.

Well, there you have it, a thanks to all those who helped, a glimpse at the publication process, and a caveat about what you'll be reading. Thanks for the opportunity.

Francis (Skip) Fennell
Western Maryland College

THANKS FROM THE MANAGING EDITOR

All twenty-six manuscripts that were submitted for publication in the RCDPM Research Monograph were circulated to a panel of referees who reviewed the manuscripts and made recommendations regarding their publication. Every manuscript was read by a minimum of three different people. The work of the RCDPM referees has contributed to the quality of this Monograph and their efforts are much appreciated. Thanks so much to all of you.

The people listed here have served as referees within the past year. If the name of any referee has been omitted, it is an error, not a lack of appreciation.

Francis (Skip) Fennell
Managing Editor

Douglas B. Aichele
Robert Ashlock
Ian D. Beattie
Nancy Bley
Raymond Brie
James V. Bruni
Grace M. Burton
Leroy G. Callahan
Helen Cheek
Douglas Edge
Jon M. Engelhardt
Cathey Funk
Linda O. Gemmill
E. Glenadine Gibb
Neil Hall
James W. Heddens
Jane M. Hill
Boyd Holtan
Martha Hopkins
Mary Ellen Hynes

Michael Hynes
Martin Johnson
Carol Novillis Larson
Beth Lazerick
Nancy Levine
John D. McGregor
Theodora R. Meister
Douglas Owens
Sandra Powers
Nina Ronshausen
Barbara Sadowski
Janet K. Scheer
G. Bradley Seager
Dora H. Skypek
Bill Speer
Marilyn N. Suydam
Cecil R. Trueblood
Robert Underhill
A. Edward Uprichard
Nancy M. Wilson

Contents

Preface.	i
The Mathematics Clinic Model: Parents as Clinicians and Their Effect on Achievement, Attitude, and Self-Concept Lee Ann Wise and Janet K. Scheer.	1
Heterogeneous vs. Homogeneous Item Arrangements in Mathematics Testing Jon M. Engelhardt, Linda M. Oldaker, Leonard S. Cahen	10
Testing Mathematical Problem Solving Thomas A. Romberg, Kevin F. Collis, Murad E. Jurdak	20
A Model and Sequence for Diagnostic Instruction in Problem Solving Francis (Skip) Fennell.	33
The Validation of a Criterion-Referenced Test of Rational Number Concepts and Skills Barbara R. Sadowski	41
An Exploration: White and Clark's Test of Inclusion Applied to Ex Post Facto Learning Data Bob Underhill	50
Brain Hemispheric Considerations for a Developmental Approach to Diagnosing Mathematics Concepts for Children Janice L. Flake	65
Doing It Their Way: Some Child-Methods in Mathematics Lesley R. Booth, Kathleen Hart.	80
Diagnostic and Prescriptive Instruction Applied to Teaching Mathematically Talented Students Kevin G. Bartkovich	85
Linguistic Aspects of <u>More/Less</u> and Dimensional Adjectives Relevant to Diagnosis and Instruction Carol Novillis Larson	95
Improving Computation--A Look at Estimation and Approximation as Significant Elements in Achieving Understanding and Correctness Douglas H. Crawford	107
Strategies for Diagnosing and Enhancing Learning of Place-Value Concepts Bob Underhill	115

The Mathematics Clinic Model:
Parents as Clinicians and Their Effect on
Achievement, Attitude, and Self-Concept

Lee Ann Wise and Janet K. Scheer
Southern Illinois University

In recent years, there has been an increased emphasis within the schools towards humanistic education. Saylor (1975) defined humanistic education as

people working together cooperatively to enhance human personality in directions satisfying both to the individual personally and to the other members of the social group (p. 16).

Saylor further stated that one of the main components of humanistic education is that the role of the family (i.e., parents, siblings, and an extended family) must be fully recognized. In the past, professional educators have been highly remiss in efforts to work cooperatively with the family to promote maximum development of a child's potential.

Changes in parent participation are leading to a level of involvement where parents and teachers are considered partners in the education process; both share their knowledge on child development, education, and individual children. Many opportunities should be provided by the school to involve the parents in all aspects of the education of their children. One way to accomplish this would be for school districts to develop programs aimed at sharing the information that parents and teachers have about children and their needs and to plan ways in which to cater to those needs. Bettering the communication between parents and schools and, hopefully, bettering the student's education is what parental involvement is all about (Office of Education, 1978).

Parental involvement in education has been shown to have a positive effect on children's learning. Douglas (1964) cited his study of approximately 5,000 British children born in a single week in March, 1946. Parents of these children were judged to show a high, fair, or low level of interest, based on (a) teacher comments on the level of parent interest in the child's school work, and (b) records of the number of times each parent visited the school to discuss the child's progress. At both 8 and 11 years of age, the highest average mental test scores were made by children whose parents were most interested, with a relationship found between test scores and parental interest within each social class. Douglas concluded that:

The influence of the level of the parent's interest on test performance is greater than that of any of the other three factors: size of family, standard of home, and academic record of the school (p. 87).

In the project sponsored by Mobilization for Youth (Lazar & Chapman, 1972), first-grade children whose parents were trained one hour per week to read to their children scored higher on nine different reading tests than did either matched children who received two hours of remediation per week from professionals or a control group receiving no treatment.

Research pertinent to the social environment of culturally disadvantaged children as it affects achievement motivation has shown that continuity between the home environment and the school appears to influence strongly a child's achievement in school. Willmon (1969) noted that parental involvement in an educational program appears to be a factor in future academic motivation and that programs concerned with upgrading student achievement should involve and educate parents.

Since teachers are professional educators trained and paid to teach, one is tempted to ask why parents should be involved in the educational process. The reason is that, as research has shown time and again (Cullen, 1969; Douglas, 1964; Hess, 1969; Schaefer, 1972; Rupp, 1969), parents are the primary models for their children's behavior. Therefore, if parents are involved and concerned with educational matters, their children will be too. Many believe, in fact, that parental involvement is crucial to fostering a willingness to learn and guaranteeing success in the undertaking (Koppman, 1978).

Due to the positive relationship between parental involvement and student achievement in school in the above-cited studies, this study was designed to investigate whether parent participation in supplemental diagnostic/prescriptive mathematics activities is associated with changes in their own children's mathematics achievement, attitude toward mathematics, and self-concept, and to compare those results to those of students attending the supplemental diagnostic/prescriptive clinic at Southern Illinois University.

The Problem

Parents are the first and most continuous teachers of their children both at home and in the classroom. Traditionally, many parents have felt that they have no meaningful roles to play in the school setting. The resulting alienation has severely limited parent usefulness as resource persons and partners in the development of their children (Follins, 1979). This study attempted to alleviate this practical problem by involving parents in their children's education and by training them to provide supplemental diagnostic/prescriptive mathematics learning experiences for their children.

The review of literature revealed that there exists a paucity of research concerned with parents providing supplemental diagnostic/prescriptive mathematics activities for their children. The current study provided additional information in this area by utilizing the Southern Illinois University Diagnostic Math Clinic model to train parents to work with their own children who have been perceived to have difficulty learning mathematics. The study investigated possible positive changes in a child's mathematics achievement, attitude toward mathematics, and self-concept as a result of working after school with his/her own parent who was utilizing diagnostic/prescriptive techniques.

Clinicians in the Southern Illinois University Diagnostic Math Clinic are in-service and pre-service teachers who work one-to-one with children (a) administering general diagnostic mathematics tests, cognitive style tests, and specific diagnostic tests; (b) analyzing the results of these tests; (c) hypothesizing the cause of the difficulties; and (d) correcting the deficiencies. Clients in grades 2 through 12 are recommended to the clinic by parents or teachers who observe a deficiency in the children's mathematics. Recognizing that they are unable to deal with every child's problem on an individual basis, teachers welcome the opportunity that the Math Clinic offers.

Because of its reputation as well as the fact that it is the only service of its type in the southern Illinois area, the Diagnostic Math Clinic has a long waiting list and must operate on a first-come-first served basis.

Quality instruction is as important as being able to offer help to all who require it. The researchers perceived a need to extend the services of the Math Clinic. According to the literature reviewed, parents could be a viable alternative and, therefore, were utilized in this study.

Method

Subjects

In August, 1981, the researchers secured permission from the superintendents of Giant City School and Unity Point School, both in Carbondale, Illinois, to distribute registration forms for the newly established Parent-Student Mathematics program offered at Southern Illinois University. Parents and children who agreed to participate consisted of students from the above mentioned schools as well as students who were on the waiting list for the Southern Illinois University Diagnostic Math Clinic.

The sample consisted of 30 children from a variety of backgrounds. They ranged in age from 8 to 13 years old and were enrolled in grades 3 through 8 in southern Illinois area schools. Each group contained 11 males and 4 females. The experimental group consisted of 15 students (average age 10.6 years and average grade level 5.1 years) who participated in the Parent-Student Mathematics program at Southern Illinois University. This program was offered twice. The first session was held during the Fall of 1981, and the second session was held in the Spring of 1982. Ten students participated in the Parent-Student Mathematics Program during the Fall session, and five students participated during the Spring session. Each parent worked with his/her own child during the course of the program.

The control group, which also consisted of 15 students, was matched to the students in the treatment group by age, grade level, and gender. Average age of this group was 10.7 years and average grade level was 5.4. The students in the control group were participants in the Diagnostic Math Clinic at Southern Illinois University during the Fall of 1981 or Spring of 1982.

Design

Both the experimental and the control group were pretested before their sessions began and posttested at the conclusion of their supplemental diagnostic/prescriptive sessions. Pretesting and posttesting consisted of a battery of three tests: Key Math Diagnostic Arithmetic Test (Connolly, Nachtman, & Pritchett, 1976), Revised Math Attitude Scale (Dutton, 1962), and the Self-Appraisal Scale (Davidson & Greenberg, n.d.). These tests measure mathematics achievement, attitude toward mathematics, and self-concept, respectively. All tests were individually administered. Testing took place at Southern Illinois University.

Key Math Diagnostic Arithmetic Test has 14 subtests organized into three sections: Content, Operations, and Applications.

<u>Content</u>	<u>Operations</u>	<u>Applications</u>
Numeration	Addition	Word Problems
Fractions	Subtraction	Missing Elements
Geometry & & Symbols	Multiplication	Money
	Division	Measurement
	Mental Computation	Time
	Numerical Reasoning	

The test provides four levels of diagnostic information: 1) total test performance, 2) area performance, 3) subtest performance, and 4) item performance. This study used the data provided by level 1 only as information gained from levels 2, 3 and 4 would be irrelevant to the hypothesis. The information gained from level 3 indicated specific areas which needed work but the information was irrelevant in terms of the results of the statistical analysis.

In the experimental group, four students worked on place value and addition, four on place value and subtraction, two on multiplication, two on division, and three on fractions. In the control group, two students worked on place value and addition, eight on place value and subtraction, two on multiplication, two on division, and one on fractions. However, only the scores for total test performance on the Key Math were used in the data analysis.

Because the parents were not trained to administer tests, test administrators were utilized to pretest and posttest the experimental group. These administrators had all participated in a 16-week 3-credit graduate course at Southern Illinois University entitled "Diagnostic and Corrective Techniques of Mathematics." The tests were administered to the control group by their clinicians.

Training

Because the researchers expected parents to utilize diagnostic/prescriptive techniques when working with their children, training sessions were an integral part of this study. Parents received six hours of training (three 2-hour sessions) on the use of manipulatives and mathematics games, making materials and games, reading and interpreting achievement, attitude, and self-concept test scores, teaching strategies, utilizing the concrete-iconic-abstract learning model, writing lesson plans, questioning techniques, use of positive reinforcement, verbal and nonverbal behavior, and record keeping. The main goal of the sessions was to train parents to become qualified supplemental prescriptive teachers of their own children in the area of mathematics.

Clinicians who worked with the children in the control group (participants in the Diagnostic Math Clinic) received in-depth training in each of the above mentioned areas while they were students in a 16-week 3-credit Southern Illinois University course, "Diagnostic and Corrective Techniques in Mathematics" during the semester prior to their work in the Clinic. During this time, they also learned to administer achievement, attitude, self-concept, and other cognitive style tests, and to write their own diagnostic tests.

Remedial Sessions

Remediation for both the control and experimental groups consisted of 16 one-hour sessions (2 one-hour sessions each week). In the experimental group, a parent worked with his/her own child; in the control group a clinician was assigned to one child for the entire semester. Every session was conducted on a one-to-one basis and was geared toward each individual child's attitude, self-concept, achievement, learning styles, and interests. For both groups, one of the hour sessions each week was conducted at Southern Illinois University where the parents and the clinicians were under close supervision of the researchers. The second hour session each week was held at a specified time and place decided upon by the parent and child or the clinician and child. There was no researcher supervision during the second session.

Instruction during all sessions utilized a positive approach where the child's strengths were used in treating areas of weakness. Instruction moved from concrete to iconic to abstract in order to promote understanding of mathematical concepts. Home-made and commercial manipulatives and games, along with the Mathputer, Little Professor, DataMan, and Speak & Math, were utilized during remedial sessions. The researchers met once a week with both parents as a group and clinicians as a group to discuss successful experiences and problems encountered during the remedial sessions. Suggestions and recommendations were made by the researchers.

Results

This study investigated the effect of parents utilizing supplemental diagnostic/prescriptive techniques on their children who were perceived to have difficulties with mathematics. Table 1 shows the pretest and posttest raw score means on the three dependent variables (mathematics achievement, attitude toward mathematics, and self-concept) for the experimental and the control group. Within-group change was evaluated using dependent *t*-tests, the results of which are also shown in Table 1. For mathematics achievement, significant mean increases were found for both the experimental and the control group. No significant difference was found for either group on attitudes toward mathematics or self-concept.

Posttest scores for the experimental and control group were compared using an analysis of covariance with pretest scores as the covariate, as recommended by Campbell and Stanley (1963). Other possible covariates such as age, grade level, and gender were not used since the groups had already been matched by those variables. Table 2 shows the adjusted posttest means for each group along with the corresponding *F*-values for differences between the means. Nonsignificant differences were found for each of the three dependent variables.

Conclusions

The fact that there were no significant differences between posttest scores on the Key Math Diagnostic Arithmetic Test between the control group and the experimental group would indicate that the achievement of children who are taught mathematics in a supplemental setting by their parents, who have utilized diagnostic/prescriptive techniques, is as great as those children who have participated in the Southern Illinois University Diagnostic Math Clinic. Significant differences did not appear between posttest scores on the Revised Math Attitude Scale and the Self-Appraisal Scale. Past research (Scheer, 1977) shows that changes in attitudes and self-concepts often require more time than do changes in achievement.

Table 1

Dependent t -Values for Pretest vs. Posttest
Mean Scores on Mathematics Achievement,
Attitude Toward Mathematics, and Self-Concept

Variable	Pretest Mean Score	Posttest Mean Score	Dependent t -Value
Mathematics Achievement			
Experimental	135.93	148.47	4.93*
Control	127.33	139.93	4.49*
Attitude Toward Mathematics			
Experimental	0.23	0.55	1.68
Control	0.02	0.44	1.85
Self-Concept			
Experimental	57.33	57.87	0.38
Control	53.80	55.93	2.01

*Significance at $p < .001$

Table 2

Analyses of Covariance for Experimental vs.
Control Group on Mathematics Achievement,
Attitude Toward Mathematics, and Self-Concept

Variable	Experimental Group Adjusted Posttest Mean Scores	Control Group Adjusted Posttest Mean Scores	F
Mathematics Achievement	148.45	139.92	0.05
Attitude Toward Mathematics	0.55	0.44	0.05
Self-Concept	57.89	55.96	0.02

The results of this study seem to have major implications for the education of children and for the academic community. Because of the economic crunch, many school districts are eliminating programs which help children with mathematics deficiencies. Trained parents may be utilized by the schools to work with such children. It would be necessary, then, for universities to train parents to utilize the diagnostic/prescriptive techniques in mathematics. Potential and in-service teachers would also need to be made aware of the possibility of utilizing parents as a resource in their classrooms.

Teachers frequently complain that children do not retain their mathematics during the summer. If parents are trained to use diagnostic/prescriptive techniques, they could effectively work with their children during extended vacations from school. This kind of at-home activity could quite possibly eliminate the need for extensive review and reteaching at the beginning of every school year.

The study described in this paper is only a small step in determining the effectiveness of parents conducting supplemental diagnostic/prescriptive mathematics sessions with their own children. Further research is needed to investigate (a) the effects of parents working with children other than their own, (b) parent attitudes and self-concepts, (c) sibling tutoring, (d) the involvement of senior citizens (especially retired teachers) as tutors, and (e) the effectiveness of parents aiding a teacher in the classroom rather than after school. Another extension of this study would be to include a second control group to the design which consists of parents and their children who are willing to participate but receive no treatment. This addition would help control for 'parent willingness to participate' being an extraneous variable that may have affected the outcome of the program. The results of this study may warrant establishing a structured model that could be used to train parents to utilize diagnostic/prescriptive techniques to improve the achievement, attitude, and self-concept of children who exhibit difficulty with mathematics. Such a structured training model could then be used by universities and school districts who work with parents.

References

- Campbell, D. T., and Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.
- Connolly, A.J., Nachtman, W. and Pritchett, E.M. Key math diagnostic arithmetic test. Circle Pines, Minn.: American Guidance Service, Inc., 1976.
- Cullen, K. School and family. Dublin: Gill and Macmillan, 1969.
- Davidson, H., and Greenberg, J. Self-appraisal scale. New York: The City University of New York, n.d.
- Douglas, J.W.B. The home and the school: A study of ability and attainment in the primary school. New York: Ballantine, 1964.
- Dutton, W.H. Attitude change of prospective elementary school teachers toward arithmetic. The Arithmetic Teacher, 1962, 9, 421.
- Follins, J.R. Diagnostic targeted assistance (DTA): A process to assess qualitative and quantitative involvement of parents in the Clark College home-school partnership model (HSPM). San Francisco, Calif.: American Educational Research Association's Annual Meeting, April 8-12, 1979 (ERIC Document Reproduction Service No. ED 171 754)
- Hess, R.D. Parental behavior and children's school achievement: Implications for Head Start. In E. Grotbert (Ed.), Critical issues in research related to disadvantaged children. Princeton, New Jersey: Educational Testing Service, 1969.
- Koppman, P.S. Taking learning home. St. Ann, MO: Central Midwestern Regional Educational Lab, June 1978. (ERIC Document Reproduction Service No. ED 157 018)
- Lazar, J.B., and Chapman, J.E. A review of the present status and future research needs of programs to develop parenting skills. Washington, DC: Social Research Group, George Washington University, April 1972.
- Office of Education (OHEW). Parents as models for reading. St. Ann, MO: Central Midwestern Regional Education Lab., June 1978. (ERIC Document Reproduction Service No. ED 157 018)
- Rupp, J.C.C. Helping the child cope with school: A study of the importance of parent-child relationships with regard to elementary school success. Gronigen: Walters-Noordhoof, 1969.
- Saylor, J.G. Humanistic education: The minimum essentials. In Glen Haas (Ed.), Curriculum planning: A new approach (3rd Ed.). Boston: Allyn & Bacon, Inc., 1980.

- Schaefer, E.S. Parents as educators: Evidence from cross-sectional, longitudinal, and intervention research. Young Children, 1972, 27, 227-239.
- Scheer, J.K. Effects of supplemental diagnostic/prescriptive teaching on mathematics achievement, attitude, and self-concept. (Doctoral dissertation, Arizona State University, 1977). Dissertation Abstracts International, 38A: 681; August 1977. (77-16, 804)
- Willmon, B. Parent participation as a factor in the effectiveness of Head Start programs. Journal of Educational Research, 1969, 62, 406-410.

Heterogeneous Vs. Homogeneous Item Arrangements in Mathematics Testing

Jon M. Engelhardt
Linda M. Oldaker
Leonard S. Cahen
Arizona State University

Mathematics tests can take on various formats. Items that sample a specific domain of knowledge may be presented in at least two ways: (1) many items over a given objective/skill are presented in groups of clusters (i.e., homogeneous groupings), and (2) only a few items are presented and those items are not necessarily juxtaposed (heterogeneous groupings). The question arises as to which of these formats provides a better estimate of subjects' domain performance. These two different presentations may yield different reflections of subjects' understanding, depending upon students' achievement level.

The investigators reasoned that when items are arranged so that consecutive items differ--as compared to a clustered format--the subjects expend more effort in deciding appropriate rules. In so doing, the subjects consume more time, which may possibly induce greater fatigue and increase the probability of confusion and different approaches to similar problems. All of this is likely to lead to lower performances on the test. This may be especially true for students of low mathematics ability, who generally are more frustrated, have lower motivation, deal less well with a lack of organization, are more easily confused, and respond impulsively (e.g., Schulz, 1972; Lesh, 1979).

Performance differences resulting from heterogeneous and homogeneous item arrangements need study. As already noted, some mathematics tests use heterogeneous item arrangements, while others, such as criterion-referenced tests assessing a single domain, use homogeneous arrangements. If these organizations do in fact affect performance, particularly for low achievers, then they may lead to different estimates of a pupil's mastery/proficiency level.

The Study

In this investigation, mathematics performances on heterogeneous and homogeneous test-item arrangements were compared. Specifically, the purposes of the study were to determine if test-item arrangement (heterogeneous vs. homogeneous) affected students' performance on a mathematics test, both generally and differentially depending on students' ability levels.

Method

Subjects

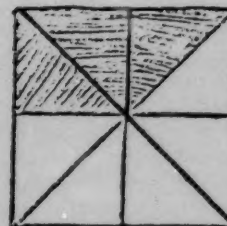
The subjects for this investigation were 45 fifth-grade students, representing all of two classes from a private school in a suburb of Phoenix, Arizona. Neither class represented a homogeneous ability group in mathematics, reading, or any other subject matter area.

Materials

Two specially constructed mathematics tests were used in this study. Both tests consisted of free-response items; one having a homogeneous format and the other, a heterogeneous format. Each test sampled the same three knowledge domains: (a) concept of a proper fraction, (b) multidigit subtraction with regrouping (no zeros), and (c) two-digit by three-digit multiplication with regrouping. Eighteen items of roughly equivalent difficulty were presented in individual test booklets, generally two items per 8 1/2" by 5 1/2" page. A sample item for each of the three domains is presented in Figure 1.

7. What fractional part of this region is shaded?

Proper fractions



Multidigit subtraction (Regrouping)

$$\begin{array}{r} 6. \quad 738 \\ - 89 \\ \hline \end{array}$$

Two-digit by three-digit multiplication (Regrouping)

$$\begin{array}{r} 11. \quad 289 \\ \times 42 \\ \hline \end{array}$$

Figure 1. Sample items from the three knowledge domains.

The homogeneous test (Hm) consisted of 54 items clustered in three groups of eighteen: fractions, subtraction, and multiplication, in that order. All subjects received the same version of this test. The heterogeneous test (Ht) likewise consisted of 54 items, but with only three items representing each domain. These nine items were distributed throughout the test, each subject receiving a different set of nine items. Items were selected in groups of three by matrix sampling (Lord & Novick, 1968) using the 18 items from the homogeneous test as an item pool for each domain. Thus collectively the heterogeneous tests contained the same items as the homogeneous test. The additional 45 filler test items came from other knowledge domains in mathematics. They were created to resemble items on a standardized mathematics achievement test (Madden, Gardner, Rudman, Karlsen & Merwin, 1972). If items were originally multiple choice, they were rewritten as free response items.

Besides the test material, worksheets for each of the three knowledge domains were assembled. Some of these were pages of problems for drill and practice; others were games and puzzles. These were to be used by classroom teachers to review subjects over the three knowledge domains (so that item arrangement was the main variable being examined in the experiment).

Procedure

All testing was conducted during one week in the spring of 1981. On Monday and Tuesday subjects received structured review using the worksheets. This was done to control for subjects' recency of exposure to the ideas tested. On each of the two days following, subjects received one of the two tests. The students were told they were going to take a two-part test that would indicate their progress in mathematics. Half of each class was randomly selected to receive the homogeneous test on the first day and the heterogeneous test on the second day. The other half received the two tests in the reverse order. Students were allowed up to 50 minutes to complete each day's testing. All reviewing and testing was conducted in the classroom by the classroom teacher.

Results

Eight students were removed from the analysis because they missed either or both tests. These absences appeared to bear little relationship to class membership or test performance. Only five students found the 50 minutes insufficient to complete the daily testing, and this on only the heterogeneous testing. Since the incomplete testing involved but one problem from the domains being tested, no further data was discarded from analysis.

For the remaining 37 subjects, means and standard deviations for performance on the two test formats were calculated. The results are displayed in Table 1. Note that even though the investigators expected the fraction items to be relatively easy, students did exceptionally well; only two students missed more than an occasional fraction item. Error analysis revealed inappropriate inversions (e.g., $3/2$ for $2/3$), giving the appropriate fraction for the unshaded portion of the region, or writing the fraction as a ratio of shaded to unshaded parts (e.g., $1/5$ as $1/4$). Similarly, subtraction performance was quite good, and most of the subtraction errors were made by four students. Errors resulted primarily from difficulties with the regrouping process and

Table 1
Means and Standard Deviations for Test Performance
($n = 37$)

Subtest	Versions			
	Homogeneous (18 items)		Heterogeneous (3 items)	
	M	SD	M	SD
Fraction	16.7	2.8	2.8	0.5
Subtraction	15.2	3.9	2.6	0.9
Multiplication	10.2	5.4	1.8	0.9
Total	42.1	8.7	7.1	1.6

occasional use of incorrect subtraction facts. The multiplication subtest led to much more variable performances. Errors on multiplication items tended mostly to be basic fact errors and a variety of difficulties with the algorithm. Errors seemed to be concentrated in the performance of sixteen students.

In analyzing the data, for each of the nine items appearing on both the homogeneous and heterogeneous tests (three items per knowledge domain), subjects were assigned an ordered pair of numbers. The first number indicated performance on the homogeneous (18 item) version and the second number the heterogeneous version. A "0" indicated an incorrect response and a "1" a correct response. Figure 2 shows a few samples of individual performances.

As may be noted, there are only four possible ordered pairings: 00, 01, 10, 11. Examining the distribution of these pairings should lend support to or refute the contentions made in this investigation. The pairings 00 and 11 indicate an individual got an item appearing on both tests incorrect and correct, respectively. Although one might argue that the pairing 11 fails to support the notion of heterogeneous item organization leading to lower performance, the pairing 00 adds little evidence pro or con. The pairing 10 is clearly support for the contention, while the pairing 01 refutes it. Thus if the main contention of this investigation is true, one would expect that of those individuals failing to get items correct or incorrect on both tests, the predominance of students would tend to perform worse on the heterogeneous version than on the homogeneous version (i.e., the pairing 10), particularly low-achieving students.

Figure 3 shows in cross-tabulation format paired item performance for the three domains, separately and combined. (There were only 110 item pairings per domain because one item in each domain was accidentally given twice to the

Subject #4			Subject #8			Subject #10		
Item	Hm	Ht	Item	Hm	Ht	Item	Hm	Ht
11	1	1	3	1	1	9	1	1
Fr. 14	1	1	Fr. 4	1	1	Fr. 10	1	1
16	1	1	15	1	1	18	1	1
29	1	1	21	1	1	20	1	1
Subt. 32	1	1	Subt. 22	1	1	Subt. 29	1	1
34	1	1	33	1	1	36	1	1
47	1	0	39	0	0	44	1	0
Mult. 50	1	1	Mult. 40	0	0	Mult. 45	0	1
52	0	0	51	0	0	47	1	1

Subject #19			Subject #27		
Item	Hm	Ht	Item	Hm	Ht
1	1	1	2	1	1
Fr. 2	1	0	Fr. 9	1	1
3	0	0	12	1	1
20	1	1	20	1	0
Subt. 21	1	1	Subt. 23	0	0
23	1	1	27	0	0
37	1	0	38	1	0
Mult. 38	1	0	Mult. 41	0	1
42	1	0	54	0	0

Figure 2. Sample individual performances for more items appearing on both homogeneous and heterogeneous versions.

COMBINED DOMAIN PERFORMANCE

		Ht Version	
		0	1
Hm Version	0	34	32
	1	34	230

SEPARATE DOMAIN PERFORMANCE

FRACTIONS

		Ht Version	
		0	1
Hm Version	0	6	1
	1	3	100

SUBTRACTION

		Ht Version	
		0	1
Hm Version	0	7	6
	1	9	88

MULTIPLICATION

		Ht Version	
		0	1
Hm Version	0	21	25
	1	22	42

Figure 3. Paired item performance for the three domains, combined and separated.

same person, a different person for each domain.) In examining the combined performance, most students tended to get items in both versions correct (70% of the items). Of those students failing to get an item correct on one or the other version (01 and 10), the same proportion of items (10%) were correct on both versions. Stated another way, of those items responded to incorrectly on the heterogeneous version, half were correct on the homogeneous version, but similarly half the items incorrect on the homogeneous version were correctly answered on the heterogeneous version.

When paired-item performance for each knowledge domain is examined separately, the same general trend appears. In the case of multiplication, where performance was generally lower and more variable, a greater proportion of item pairings showed an incorrect response; however, the proportion of 01 and 10 pairings remain about equal.

Using performance above and below mean performance on the homogeneous test to define high and low achievers, a similar inspection of item pairings was made. Except for greater and lesser concentrations of pairing 11 for high and low achievers respectively, the same trend was observed for both groups. Likewise, further examinations considering test order and teacher factors led to identical results.

As presented in Figure 2, each subject's performance was examined for idiosyncratic differences. As may be noted from this sample of subjects' performances, no clear pattern of support was found. It was expected that a number of subjects would perform like Subject #19 (relatively consistent presence of pairing 10), particularly among low achievers, but this was the only subject like that. Most showed mixed performance (both pairings 01 and 10), and two even showed somewhat consistent 01 performance.

In light of the results presented thus far, no further analyses were pursued.

Discussion

The researchers expected that, particularly for low achievers, a test item presentation clustering a relatively large number of items for a given knowledge domain would lead to higher performance than a presentation where only a few items were used and where those were scattered among items representing other domains. The results of this study indicated no support for that expectation (nor, incidentally, was there support for the reverse relationship). If any conclusion seems warranted, it is that in the vast majority of the cases students consistently respond correctly or incorrectly (in general, 80% of the items), regardless of whether an item is one of the three or one of eighteen and regardless of whether an item is juxtaposed with similar items or scattered among dissimilar items. The results may even imply that three free-response items are as adequate as eighteen in sampling a knowledge domain. (In fact, a correlated t-test comparing the proportion of correct responses on the 3-item and 18-item tests yielded no statistically convincing evidence of difference in performance. This was true for each knowledge domain, separately and combined, and for high and low achievers. (See Table 2.)

Table 2
Comparison of Proportion of Correct Responses on
Homogeneous and Heterogeneous Versions

Domains	Students	<u>n</u>	Correlated <u>t</u>	<u>p</u>
Fractions	All	37	0.47	0.64
	Low achievers	17	-0.78	0.40
	High achievers	20	1.47	0.16
Subtraction	All	37	-0.29	0.78
	Low achievers	17	-0.11	0.91
	High achievers	20	-0.34	0.74
Multiplication	All	37	-0.76	0.45
	Low achievers	17	-1.44	0.17
	High achievers	20	0.76	0.46
Combined	All	37	-0.60	0.55
	Low achievers	17	-1.26	0.22
	High achievers	20	0.80	0.44

In attempting to account for these unexpected results, one might hypothesize that, especially on the multiplication items where the students encountered more difficulty and likely worked more slowly, subjects became bored or tired by the eighteen items, thus leading to as similar a decrement in performance as that caused by the possible distraction, fatigue, and frustration of constantly switching knowledge domains on the tests with three items per domain. If this were true, one would expect a general deterioration of performance toward the end of each 18-item cluster. Inspection of the data suggests, however, that, while this was the pattern for a few students (about eight for multiplication), it was generally not the case. Most students performed rather consistently throughout the eighteen clustered items.

Although the subjects in this investigation were from a private school, the fact that each subject was her or his own control makes it hard to claim sampling limitations as a confounding influence. The relative lack of difficulty with the fraction and subtraction items may have made differences less obvious, but the multiplication items, by themselves, support the conclusion (of consistent response despite arrangement) in this investigation.

No alternate explanation for the results of this study is apparent. Before any wholesale application of the findings of this study is appropriate, however, replication with a larger and more diverse sample, including varying age levels, is recommended. Clearly, though, the use of limited numbers of disbursed items to imply mastery of a domain of knowledge appears to be no better or worse than the use of large numbers of juxtaposed items.

References

- Lesh, R. Mathematical learning disabilities: Considerations, for identification, diagnosis, remediation. Manuscript, 1979. (available from R. Lesh, Northwestern University.)
- Lord, F. M., & Novick, M. R. Statistical Theories of Mental Test Scores. Reading, Mass.: Addison-Wesley, 1968.
- Madden, R., Gardner, E., Rudman, H., Karlson, B., & Merwin, J. Stanford Achievement Test. New York: Harcourt, Brace and Jovanovich, 1972.
- Schulz, R. W. Characteristics and needs of the slow learner. In W. C. Lowry, The Slow Learner in Mathematics. Washington, D.C.: National Council of Teachers of Mathematics, 1972.

Testing Mathematical Problem Solving

Thomas A. Romberg, University of Wisconsin-Madison, USA
 Kevin F. Collis, University of Tasmania, Hobart, Tasmania, Australia
 Murad E. Jurdak, American University of Beirut, Beirut, Lebanon

Introduction

The purpose of this paper is to outline the development of a set of super-items designed to assess mathematical problem solving and to report on the content and construct validity of the items. A "super-item" is a set of test questions based on a common situation or stem. In this project, a pool of mathematical situations and a set of items for each situation were designed to provide information about students' qualitatively different levels of reasoning ability. The items for each situation were designed to elicit increasingly complex levels of reasoning from the students. The structure was based on Biggs and Collis' SOLO taxonomy (1982). The items were prepared to be administered to students of 9, 11, 13, and 17 years of age.

The project was funded by the Education Commission of the States (with funds supplied by the National Institute of Education). The resulting items are to be used in future National Assessment of Education Progress (NAEP) studies in mathematics.

Rationale

The following question was addressed in this study: Can a set of super-items be constructed so that responses will reflect "level of reasoning" used by the students with respect to mathematical situations? In some tests, items come in groups, e.g., paragraph-reading tests with several questions on each paragraph, or table-reading tests with several items on each table. The problem situations or stems, the paragraphs and tables in the example, contain considerable information. The sets of questions with the stem are called super-items (Cureton, 1965), a term chosen to emphasize that differences among respondents in comprehension of a stem may produce correlated errors of measurement between items for the same stem. For the pool of items constructed in this study, a recently developed taxonomy based on the "structure of the observed learning outcome" (SOLO) (Biggs & Collis, 1982) was used as the blueprint for the development of superitems with four questions.

The SOLO Taxonomy, General

From different starting points, several researchers (Biggs & Collis, 1982; Marton, 1981; Case, 1980; Fischer, 1980) have devised similar models for the development of intellectual functioning in children and young adults. Biggs and Collis began by giving teachers actual classroom examples of levels of cognitive development which they would recognize from their interactions with children in classroom situations. Their analysis of responses from various sources made it clear that they were, in fact, dealing with two phenomena. The first was what they chose to call the Hypothetical Cognitive Structure (HCS), and the second the Structure of the Learned Outcomes or Responses (SOLO).

The former was closely related to the existing notion of Piaget's stages of cognitive development in which each stage has its idiosyncratic mode of functioning and, as far as intellectual development is concerned, its own set of developmental tasks. The latter, on the other hand, was concerned with describing the structure of any given response as a phenomenon in its own right, that is, without the response necessarily representing a particular stage of intellectual development.

To illustrate the structure of learned responses, suppose a class had been given a lesson on the formation of rain and the teacher asks the question, "Why is the side of a mountain which faces the coast usually wetter than the other side?" The following responses may be obtained from students:

1. "Because it rains more on the coastal side."

This response merely restates the question and fails to engage any relevant information from the previous lesson. Thus, it would be classified as pre-structural.

2. "Because the sea breezes hit the coastal side first."

In this response, an answer to the question is given, but it applies only one of a number of possible pieces of relevant information (or propositions) to the problem. Thus, it is classified as uni-structural.

3. "Because the sea breezes contain water vapor and they first strike the coastal side and so it rains on them and after that there's no rain to fall on the other side."

This response answers the question and uses several appropriate facts (or propositions) in sequence to justify the conclusion. Hence, it is a multi-structural response.

4. "Because the prevailing winds are from the sea and they pick up moisture and as they meet the mountain, they're forced up and get colder, the moisture condenses, forming rain. By the time the winds cross the mountain, they are dry."

This response not only answers the question by using several facts (or propositions) but endeavors to integrate them. The introduction of an integrating principle makes this a relational response.

5. "This is likely to be true only if the prevailing winds are from the sea. When this is so, the water vapor evaporated from the sea is carried to the mountain slopes, where it rises and cools. Cooling causes the water vapor to condense and deposit. Not only is the wind now dryer, it is then carried up the mountain further, is compressed, now warms, and thus, is relatively less saturated than before. The effect is similar to the warm climates experienced on the Eastern slopes of the Rockies in Canada in winter. However, all this makes assumptions about the prevailing wind and temperature conditions; if these were altered, then the energy exchanges would differ, resulting in quite a different outcome."

This response shows an entirely new use of the available information (or proposition). Instead of immediately accepting the conclusion implicit in the question, the student points out that there are other possibilities and uses the given data to test out some of these hypotheses. This represents an extended abstract level of responding.

Construction of Mathematical Problem-Solving Items

For this project, we hypothesized that by using the SOLO framework, one ought to be able to design items such that a series of questions based on the stem would require more and more sophisticated use of the information from the stem in order to obtain a correct result. This increase in sophistication should parallel the increasing complexity of structure noted in the SOLO categories. We used the following criteria to write questions:

- Uni-structural (U) Use of one obvious piece of information coming directly from the stem.
- Multi-structural (M) Use of two or more discrete closures directly related to separate pieces of information contained in the stem.
- Relational (R) Use of two or more closures directly related to an integrated understanding of the information in the stem.
- Extended Abstract (E) Use of an abstract general principle or hypothesis which is derived from or suggested by the information in the stem.

An example of items constructed in this manner is shown in Figure 1. The stem provides information and each question that follows requires the student to reason at a different level in order to produce a correct response.

This is a machine that changes numbers. It adds the number you put in three times and then adds 2 more. So, if you put in 4, it puts out 14.



- U. If 14 is put out, what number was put in?
- M. If we put in a 3, what number will the machine put out?
- R. If we got out a 41, what number was put in?
- E. If x is the number that comes out of the machine, number y is put in, write down a formula which gives the value of y whatever the value of x .

Figure 1. Example of a superitem written to reflect SOLO taxonomy

For the sample item in Figure 1 it was intended that in order to obtain a correct answer, the student would need to process the information in the stem in at least the following ways:

- (U) Answer: 4. One piece of information used, one closure required, the information obtainable from either the last sentence in the stem or the diagram--uni-structural response level.
- (M) Answer: 17. All the information used in a sequence of discrete closures. The stem is seen as a set of instructions to be followed in order--multi-structural response level.
- (R) Answer: 13. All the information used but in addition, the student has to extract the "principle" involved in the problem well enough to be able to use it in reverse. The student needs to have an "overview" of the instructions in the stem in order to carry out the appropriate operations--relational response level.
- (E) Answer: $y = 3x + 2$. The student has to extract the abstract general principle from the information and write it in its abstract form. This involves dismissing distracting cues, perhaps forming hypotheses and testing them, and zeroing in on the relationships involved--extended abstract level.

Item Writing and Content Validity

To prepare a set of items for use by NAEP, the items were to be administrable to students of 9, 11, 13, and 17 years of age, and they were to reflect five NAEP content categories (Number and Numeration; Variables and Relationships; Size, Shape, and Position; Measurement; Statistics and Probability; and Unfamiliar). Content validity was achieved carrying out several content vs. category checks using mathematics teachers, educators, and mathematicians. Some other aspects of validity were taken into account by extensive piloting of the items with samples of the population for which the test was intended. The pilots led to several rewritings to take account of readability and comprehensibility of the item stems. Details of this development are reported in Romberg, Collis, Donovan, Buchanan, and Romberg, 1982. Since, in the last review, the indices of agreement were quite high for judgments about content and for judgments on level of reasoning, we decided the content validity of the superitems had been demonstrated.

The Construct Validity of Each Superitem-Guttman Scalability

The notion of construct validity implies that the scores on a test can be meaningfully interpreted in terms of related concepts from a psychological theory (Cronbach, 1960). When answers to each question in a superitem are coded (1 = correct, 0 = incorrect), the yield for any student is a score vector of 1's or 0's. Since the structure of the SOLO taxonomy assumes a latent hierarchical and cumulative cognitive dimension, and the response structure associated with all lower levels, in the sense that the presence of one response structure implies the presence of all lower response structures. The five expected response patterns for each of the SOLO superitems

are (0000, 1000, 1100, 1110, and 1111) for the four questions in hierarchical order, and are called Guttman true types (Guttman, 1941). Any deviation from a true type is classified as an error. A measure of the extent to which the observed response patterns belong to Guttman true types was used as one indication of the construct validity of each item.

Test Administration

To gather appropriate information on these superitems, separate group-administered tests were prepared for the 17-year-olds and for the 9-, 11-, and 13-year-olds. Separate tests were necessary because three of the most difficult items were judged appropriate only for 17-year-olds and the tests for the 17-year-olds included the questions for all four levels of reasoning: uni-structural (U), multi-structural (M), relational (R), and extended abstract (E), whereas the tests for the three lower age levels did not include the extended abstract question.

Five test forms of seven superitems each were created for each of the two age groups (17-year-olds, and 9-, 11-, and 13-year-olds). A central Wisconsin school district serving a community of 32,000 and the surrounding rural area agreed to provide a sample of approximately 300 students in each age group for the administration of the batteries. The tests were administered during September, 1981. Test booklets were randomly distributed to students. These data, along with follow-up interview data from a small sample, were used to answer six questions addressed in the study (see Romberg, Jurdak, Collis, & Buchanan, 1982, for details).

Results

To examine whether the pattern of responses were Guttman true-types, three indices were used at each age level for each superitem. First, a coefficient of reproducibility was calculated.

$$\text{coefficient of reproducibility } (r) = 1 - \frac{\text{total no. of errors}}{\text{total no. of responses}}$$

Any response pattern which is not a true Guttman type is considered an error. Thus, if there are no patterns which are considered errors, the coefficient of reproducibility is 1 and the scale is a perfect Guttman scale. If all response patterns are errors, then the coefficient is obviously zero. Second, Proctor (1970) formulated a probabilistic representation of the observed data in order to base the acceptability of Guttman method on statistical criteria of goodness of fit rather than judgment and experience. Based on maximum likelihood procedures, a misclassification parameter (p) is calculated. Third, since p is based on the predicted distribution it can be tested by chi-square techniques. A superitem was considered to be a Guttman true type if $r > .85$, $p > .15$, and χ^2 was non-significant at the .05 level.

For 17-year-olds, 31 of the 35 superitems had a coefficient of reproducibility (r) greater than .85. The probabilities of misclassification (p) were not more than 15% for 30 of the 35 superitems. χ^2 was significant at .05 with $df = 10$ for only six superitems, five of which had a high probability of misclassification. Overall, only four superitems had practical problems which

indicate they do not reasonably reflect the SOLO taxonomy, two superitems were questionable, and 29 were satisfactory.

The results of the analysis for 13-year-olds indicated that there were eight superitems for which at least one negative indicator was found. In summary, there were 27 satisfactory superitems, three that are questionable, and five that do not reflect the SOLO taxonomy levels.

For the 11-year-olds, there were 12 superitems for which at least one negative indicator was found. We judged that there were 27 satisfactory superitems, four questionable, and five which do not reflect the SOLO levels.

The results for the 9-year-olds were similar. Since only ten items had negative indicators, we decided 27 items were considered satisfactory, three questionable, and five unsatisfactory.

In summary, for the 32 items that were administered to all four age groups, 20 were satisfactory for all ages. Furthermore, when one examines the questionable and unsatisfactory items across all ages, each appears to have a content validity problem. Only two items were questionable or unsatisfactory for just one age group. For both of those, the problem for 17-year-olds was only with the E question. Thus, when one adds to the base satisfactory items, the three superitems only administered to 13-, 11-, and 9-year-olds, we get 25 satisfactory items for those age groups. For 17-year-olds, 29 superitems were satisfactory. In general, this is strong evidence that the superitem format in which terms were constructed to fit the SOLO taxonomy forms a Guttman Scale.

Other Supporting Evidence

By contrasting the p-values for each level across age levels, a consistent picture of growth can also be shown. For example, the means for U, M, R, and E scales for the 17-year-olds are shown in Table 1. The decrease in mean performance from U to R to E is consistent with the hierarchical theory. Furthermore, since 13-, 11-, and 9-year-olds took the same forms, a cross-sectional comparison indicates increase in level by age (see Figure 2). Similarly, if one were to look at individual items, the same pattern of differences was apparent. For example, in Table 2, the p-values for U, M, R, and E are shown for superitems C3 and D2 for each age group. Although there are differences between

Table 1

Scale Means for 17-Year-Olds on U, M, R, and E for Each Form

Form	n	SOLO Response Level			
		U	M	R	E
1	61	5.13	5.52	3.02	.64
2	58	6.28	4.45	3.01	1.38
3	65	5.91	5.43	2.29	.80
4	57	6.61	5.88	3.39	.93
5	62	6.48	5.79	3.29	.92

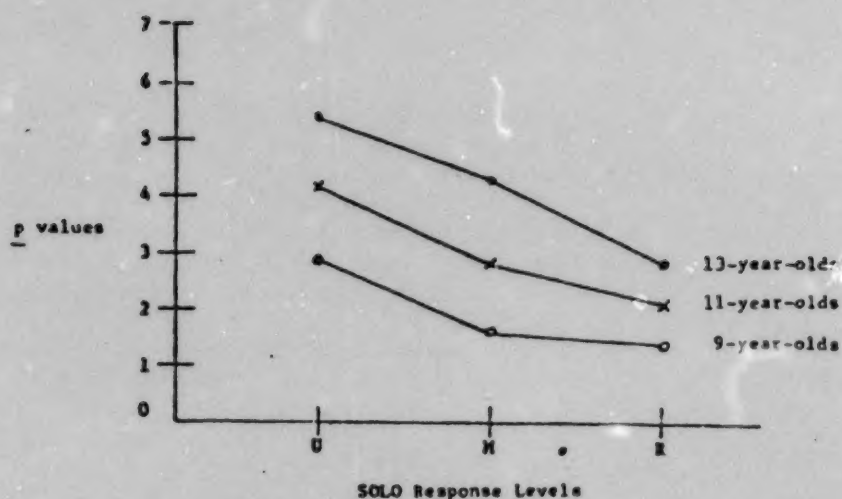


Figure 2. Profiles of p values on the U, M, and R scales for 13-, 11-, and 9-year-olds on Booklet 1 Form 2.

Table 2

p values for Superitems C3 and B2
by Level of Question for Each Age Group

Superitem	Age Level	SOLO Response Level			
		U	M	R	E
C3	17	93.0	54.4	40.4	3.5
	13	88.4	30.5	14.7	—
	11	70.3	33.8	5.4	—
	9	45.9	9.8	1.6	—
B2	17	86.2	74.1	25.9	13.3
	13	89.0	74.3	14.7	—
	11	73.2	53.7	3.7	—
	9	62.3	46.4	0.0	—

BEST COPY AVAILABLE

both forms and items, the profiles of change across age levels are consistent with the notion that there are latent cognitive levels which underlie the SOLO taxonomy and that performance is cumulative and hierarchical.

The Construct Validity--Cluster Interpretability

A second method of examining construct validity was carried out based on the following rationale. If a student's responses to a set of superitems were all Guttman true-type responses, and if the student was at a particular base stage of development (HCS), one would expect the average response pattern across several superitems to reflect that base stage of development. Furthermore, for a large number of students at any age level, one would expect that groups of students with similar response patterns for a set of items could be identified. It is plausible that the profiles of response patterns for the groups can be interpreted in terms of the SOLO taxonomy.

The profiles which would be interpretable are based on the notions of equilibrations which involve "formation instability combined with a progressive movement toward stability" (Langer, 1969, p. 93). Cognitive development is seen as "spiral" and, in particular, it is assumed that "to go forward it is necessary to go backward: the first step toward progress is regress" (Langer, 1969, p. 95). From a consideration of this notion, four suggested response profiles for students based on the SOLO superitems for two neighboring levels of performance are shown in Figure 3. The first (X) and the last (Y) steps show stability of performance at neighboring cognitive levels. The steps in between show the regression from X to Y ($X \rightarrow$) and the progression from X to Y ($\rightarrow Y$). These are seen as steps in the transition between cognitive levels.

The actual profiles were likely to be different for two reasons. First because the questions at succeeding levels are more complex, there would be an increase in the probability of making errors at higher levels. Thus, the p-values for Y will be lower than that for X. Second, since the superitems

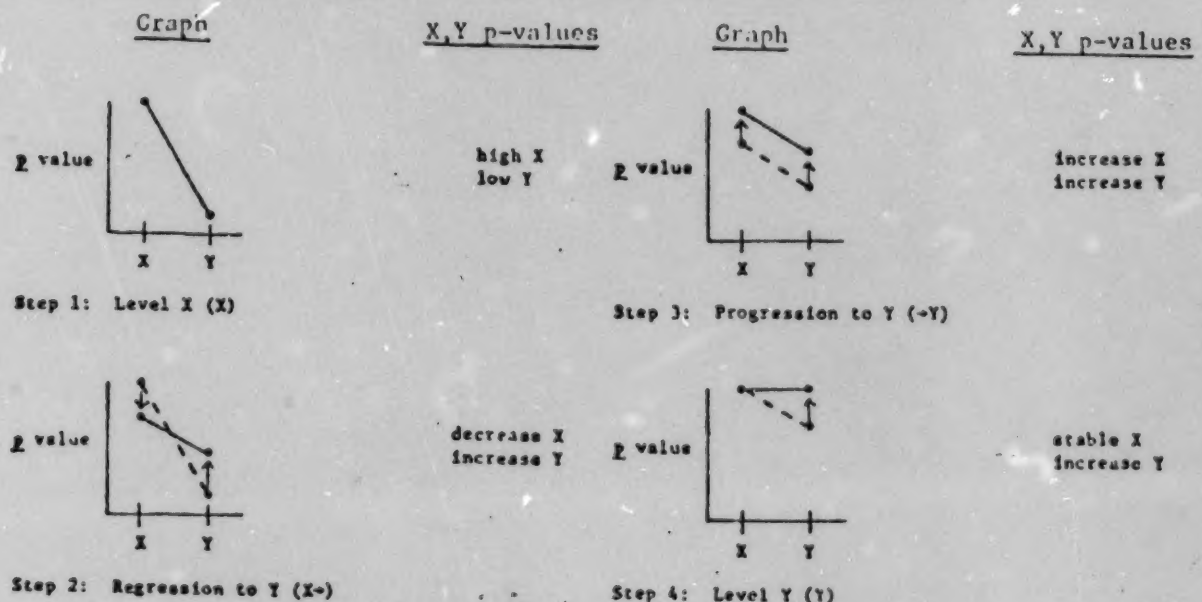


Figure 3. Four profiles of p-values for transition to neighboring cognitive levels

involve different content areas and require students to read the items, either unfamiliarity with the specific content of an item or inability to read the words would depress the patterns shown in Figure 3. If particular p-values on X and Y are similar but moderate, students would be reasoning at the Y level on those problems they understood. We decided to indicate this pattern by adding the symbol "+" to its descriptors. In summary, if a student profile for the set of superitems could be grouped with other student profiles and if the groups' average profile could be judged as similar to one of the four profiles (including depressed profiles) shown in Figure 3, then interpretability in terms of a developmental base was to be claimed.

Procedure

The maximum hierarchical clustering method (Johnson, 1970) was used to partition the students on each form and across forms into homogeneous groups based on score vectors whose four components were the aggregated scores on the four taxonomic levels of reasoning: uni-structural (U), multi-structural (M), relational (R), and extended abstract (E). Before this analysis was carried out, items which failed to reflect a Guttman scale were omitted. Clusters were found first for each form and then a sample across forms for each age group.

Results

For each of the five forms of the test given to 17-year-olds, the number of groups found varied from 5 to 7 depending upon the form. Twenty seven of 28 groups over the five forms were considered interpretable. Because of this interpretability, a random sample of 150 students was drawn from the total population. The cluster analysis of the percent correct and information for this sample is shown in Table 3. The seven groups were all interpretable. Fifty-four percent of this sample are from M to R in cognitive level, 31% are above R and 16% are below M.

For 13-year-olds, 23 of the 26 derived cluster groups were interpretable. Again, since the groups by form were interpretable, a random sample of 151 students was drawn. The cluster analysis of the profiles for this sample is shown in Table 4. Fifteen percent are at level U; another 15% are above M. Finally, 6% are between U and M and 4% are below U.

For the 11-year-old population, 28 of the 31 derived cluster groups across forms were considered interpretable. However, as one would expect, because of the lower grade level, the number of "depressed" profiles for groups increased. In general, the profiles reflect students in transition from U to M. Again, a random sample of students (154) was drawn from the population. The cluster analysis of the profiles for this sample is shown in Table 5. All seven groups are interpretable. Fifty-eight percent of the population reflect a transition from U to M ($U \rightarrow$ and $\rightarrow M$). Another 29% have reached level M (or $M \downarrow$), 11% below U, and only 9% above M.

Table 3

Percent Correct on Each Level for Cluster Groups
17-Year-Olds, Sample from All Forms

Group	n	SOLO Response Level				Label
		U	M	R	E	
1	25	99	94	79	35	-E
2	22	95	88	32	18	R+
3	41	97	92	38	07	R
4	30	90	76	33	01	M+
5	12	78	73	06	09	M
6	20	86	46	20	01	-M
7	4	45	20	00	00	-U

Table 4

Percent Correct on Each Level for Cluster Groups,
13-Year-Olds, Sample from All Forms

Group	n	SOLO Response Level			Label
		U	M	R	
1	5	.95	.94	.76	R
2	9	.92	1.00	.43	-R
3	8	.91	.67	.50	M+
4	75	.92	.77	.15	M
5	16	.50	.48	.06	M+
6	9	.72	.51	.13	U+
7	23	.94	.44	.02	U
8	6	.65	.14	.00	-U

Finally, the same procedures were followed for the 9-year-old population. In general, the patterns were more difficult to interpret. However, 23 of the 26 group profiles were considered interpretable. A random sample of students of 125 students was drawn from the population. The clusters formed from the profiles for this sample are shown in Table 6. All six groups are interpretable. Fifty-four percent of the students reflect a pattern around U ($\rightarrow$ U, U or U $\rightarrow$). Twenty-eight percent are at the P level, 18% are nearing the M level.

The consistency and interpretability of the cluster profiles across the forms indicates the stable influence of cognitive levels of development in the formation of the clusters. Furthermore, the clusters strongly support the utility of the SOLO response categories in preference to the developmental base stages in Piagetian terms. According to the theory, 17-year-old students should be at formal operational level, but most do not operate at the extended abstract level. For example, cluster group in Table 3 has the highest performance on E questions, was judged to be at the E stage of response, but contains only 16% of the students at this age. The majority of 17-year-olds operates around the relational level as seen from the size of clusters 1, 2, and 3 on the relational scale. This suggests that answering questions at the extended abstract level involves more than level of cognitive development. Similar observations are obvious at each age level. No student profiles are above the hypothesized corresponding level of cognitive development. In fact, most profiles are below the base level of development for an age group. Again, this is strong support for the SOLO levels and their utility in describing responses of students.

Table 5

Percent Correct on Each Level for Cluster Groups.
11-Year-Olds, Sample from All Forms

Group	n	SOLO Response Level			Label
		U	M	R	
1	10	.94	.91	.27	$\rightarrow$ R
2	4	.75	.46	.40	M $\rightarrow$
3	21	.99	.68	.13	M
4	13	.42	.45	.06	M $\leftarrow$
5	40	.75	.62	.09	$\rightarrow$ M
6	49	.81	.36	.02	U $\rightarrow$
7	17	.49	.18	.01	U

Table 6

Percent Correct on Each Level for Cluster Groups,
9-Year-Olds, Sample from All Forms

Group	n	SOLO Response Level			Label
		U	M	R	
1	9	.52	.52	.07	M+
2	14	.78	.57	.02	M
3	18	.80	.36	.06	U+
4	7	.71	.00	.06	U
5	42	.47	.13	.03	U
6	35	.16	.07	.01	P

In this study, we prepared a summary battery of superitems for NAEP whose questions were written to reflect the structure of learned outcomes (SOLO). We hoped to demonstrate that:

- (1) we could create a set of items which would be judged to be content valid;
- (2) the pattern of responses to each item by students would reflect the assumed latent hierarchical cognitive structure;
- (3) from the response patterns, clusters of students could be formed that would be interpretable in terms of their cognitive development.

In each case, we were remarkably successful. We have judged most of the items to be content and construct valid based on the evidence presented.

Based on these findings, we strongly recommend that researchers interested in studying level of student reasoning about complex problems (such as mathematical problem solving) follow the strategy we used of constructing superitems based on the SOLO taxonomy.

Note: This material is based upon work supported by the National Institute of Education and the Education Commission of the States under Contract No. 02-81-20321 with the Wisconsin Center for Education Research. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of ECS, NIE, or the Department of Education.

- Biggs, J. G., & Collis, K. F. Evaluating the quality of learning: The SOLO taxonomy. New York: Academic Press, 1982.
- Case, R. The underlying mechanisms of intellectual development. In J. Biggs and J. Kirby (Eds.), Instructional processes and individual differences in learning. New York: Academic Press, 1979.
- Cronbach, L. J. Essentials of psychological testing (2nd Edition). New York: Harper & Row, 1960.
- Cureton, E. E. Reliability and validity: Basic assumptions and experimental designs. Educational and Psychological Measurement, 1965, 25, 327-346.
- Fischer, Kurt W. A theory of cognitive development: The control and construction of hierarchies of skills. Psychological Review, 1980, 87(6), 477-531.
- Guttman, L. The quantification of a class of attributes: A theory and method for scale construction. In P. Horst (Ed.), The prediction of personal adjustment. New York: Social Science Research Council, 1941, 319-348.
- Johnson, S. C. Hierarchical clustering schemes. Psychometrika, 1970, 35, 241-254.
- Langer, J. Theories of development. New York: Holt, Rinehart, and Winston, 1969.
- Marton, F. Phenomenography - describing conceptions of the world around us. Instructional Science, 1981, 10, 177-200.
- Proctor, C. H. A probabilistic formulation and statistical analysis of Guttman scaling. Psychometrika, 1970, 35, 73-78.
- Romberg, T. A., Collis, K. F., Donovan, B. F., Buchanan, A. E., & Romberg, M. N. A report on the NIE/ECS Item Development Project: The development of mathematical problem-solving superitems. Madison: Wisconsin Center for Education Research, 1982.
- Romberg, T. A., Jurdak, M. E., Collis, K. F., & Buchanan, A. E. A report on the NIE/ECS Item Development Project: Construct validity of a set of mathematical superitems. Madison: Wisconsin Center for Education Research, 1982.

A Model and Sequence for Diagnostic Instruction in Problem Solving

Francis (Skip) Fennell, Western Maryland College

Problem Solving, Problem Solving, Problem Solving

There! I've said (written) it three times. Perhaps this is the feeling of many classroom teachers who recognize the importance of problem solving and yet are frustrated when it comes to finding effective strategies for presenting problem solving at the elementary school level.

For the last year or so a plethora (Howard Cosell would be proud of me!) of statements, materials, and positions have pronounced the 1980's as the decade of problem solving. Both the 1978 National Assessment of Educational Progress (NAEP) recommendations and the Priorities in School Mathematics (PRISM) survey resulted in guidelines for the treatment of problem solving in school mathematics. Both studies recommended that problem solving receive more emphasis during the current decade and be presented in a manner that stresses the use of higher-level cognitive processes. These recommendations were widely disseminated in NCTM's Agenda for Action.

Never before in the history of elementary school mathematics instruction has there been such support for teaching problem solving. Although test-construction specialists and publishers typically define problem solving in mathematics as solving word examples, to the mathematics education community a "problem" is more apt to be defined as "a situation, quantitative or otherwise, that confronts an individual or group of individuals, requires resolution, and for which the individual sees no apparent or obvious means or path to obtaining the solution," (Krulik and Rudnick, 1980, p. 3). Using this definition, problem solving becomes somewhat an individual thing. Certain word or verbal problems in textbooks may initially be problems, but after students have some practice in doing these examples many examples will be reduced to routine exercises. For many low achievers, however, we can easily understand that each mathematics encounter will truly be a problem-solving experience. The low achiever, the student whose progress is below the assigned grade level, is one whose previous experiences with problem solving have usually been negative. Since such students often make up the bulk of a clinical population, the development of an instructional sequence and of specific strategies for teaching problem solving are needed.

A Model for Teaching Problem Solving

The model for teaching problem solving described here is built on Bigge's evaluation sequence as adapted to problem-solving types (see Figure 1.). The memory level refers to recall of low-level learning activities such as counting, number facts, computational procedures, terms and vocabulary, recognition of time and money amounts, and reinforcement tasks. Explanatory understanding relates to a learning process that implies that students are able to indicate

Figure 1

A Model for
Sequencing Problem-Solving Instruction

Problem type Learning level	A. Auditory- tactile	B. Auditory	C. Pictorial- tactile	D. Pictorial	E. Word- tactile	F. Word
¹ Memory						
² Explana- tory un- derstanding						
³ Explora- tory un- derstanding						

or verify understanding through explanation. Typical oral and word problem-solving examples would, for the most part, fit into this category. Explanatory understanding may include topics like place value, problem solving, estimation and approximation activities, and a variety of related-thinking ideas. This learning category includes most actual applications of elementary school mathematics. Included in the exploratory understanding level are those activities that cause pupils to reflect and create a somewhat unique or divergent response to a learning activity. Non-routine types of problem solving, logic activities, and the use of statistics to predict outcomes represent exploratory understanding topics.

The three learning levels described may be used to represent a variety of problem-solving types. A problem-solving sequence developed for clinical use includes the following problem types:

auditory-tactile--orally stated verbal problem, which the pupil solves and validates using manipulatives.

pictorial-tactile--picture problem, which the pupil solves and validates using manipulatives.

word-tactile--a written, one-step verbal problem, which the pupil solves and validates using manipulatives.

word--a written, one or two-step verbal problem, which the pupil solves in writing or in an oral response.

Kindergarten-level pupils should have experiences that focus on auditory-tactile and auditory-word types of problems. Pupils in grades one and the first semester of the second grade should be confronted with all problem types. Students in the second semester of the second grade and continuing on through the sixth grade would benefit from encounters with problem-solving types including and extending beyond the auditory type. Based on initial clinic trials with this sequence, the actual percentages of time spent on the various problem-solving types have been determined and are presented in Figure 2.

Figure 2

Problem Solving Types, Learning Levels, and Frequency

<u>Level</u>	<u>Predominant Learning Level</u>	<u>Types of Problems</u>	<u>%</u>
K	Memory, Explanatory Understanding	Auditory-Tactile Auditory	70
1, 2 ¹	Memory, Explanatory Understanding	Auditory-Tactile Auditory Pictorial-Tactile Pictorial Word-Tactile Word	20 20 20 10 20 10
2 ² , 3	Memory, Explanatory Understanding	Auditory Pictorial-Tactile Pictorial Word-Tactile Word	20 10 10 30 30
4-6	Memory, Explanatory Understanding Exploratory Understanding	Auditory Pictorial-Tactile Pictorial Word-Tactile Word	10 10 10 10 60

Note: 2¹ refers to the first semester of second grade.

2² refers to the second semester of second grade.

Further classification involving problem types and learning levels will enable the clinician to match a particular problem type to the learning level appropriate for the pupil's progress. Most primary-grade problem types appear to be at the memory and explanatory understanding levels. Some intermediate-grade problem types involve all of the three learning levels described earlier.

Classroom Techniques for Teaching Problem Solving

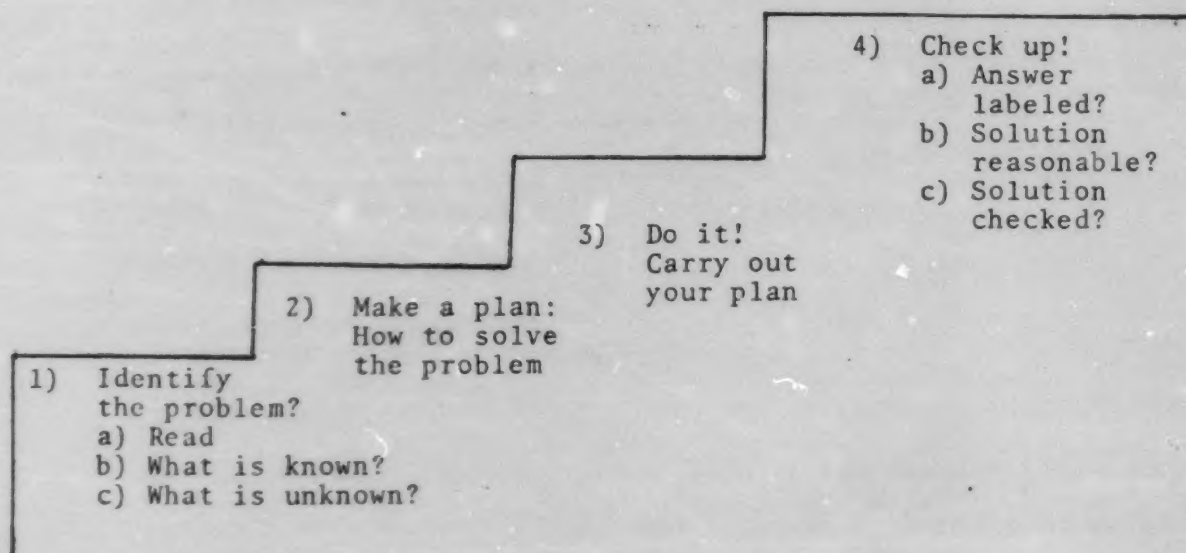
At this point it is appropriate to consider general teaching ideas that can be used with the problem solving sequence outlined in Figure 2.

Early and daily exposure. Pupils should be involved in some sort of problem solving experience each day. These daily experiences may include activities ranging from creating a "Problem of the Day" to oral discussion of textbook-related problems. Problem-solving materials should relate to mathematics concepts and ideas to which children have been or are currently being exposed.

Basic problem solving procedure. Pupils should be involved in the consistent use of procedures for analyzing and solving problems. Once pupils can be involved in the actual analysis of a problem, a problem solving model or heuristic similar to the one presented in Figure 3 may serve as an appropriate visual prompt for a systematic approach to the solution of a problem.

Figure 3

Problem Solving Staircase



Collecting and maintaining problems. Consider the following suggestions for building up a reservoir of actual problems for classroom use: Actually tape problems, varying the type of information and sequence. Also consider having the students tape their own problems to try out on their classmates. Cut up old texts and workbooks to create a classroom deck of word problems. Cut out math applications sources from newspapers, magazines, phone books, catalogs, and other resources, making sure to include pictures for those all important picture problems. Once you develop a collection of problem solving activities it would be wise to match the problems to the appropriate learning level and problem type, assuming your instructional group is well diagnosed with regard to needs in the problem-solving area.

Examples of Specific Instructional Techniques

The purpose here is to describe and relate some specific teaching techniques to the problem-solving types and learning levels outlined in the model (Figure 2).

Auditory Problem Types

Pupil discussion question. A problem is given orally to the students. The teacher then poses questions for the students to discuss and answer. Consider the following example:

Heather had 5 goldfish; she gave two to Stacey. How many goldfish does Heather have left?

Questions like the following could be discussed by the pupils:

What is the problem?
How do we solve it?
How can we show this problem using counters?
What is the answer?
Make up a problem like this one about you and a friend.

Questioning-logic Activities. A series of clues is given orally (or in writing, depending on pupil grade level) one by one, followed by a question. Two examples are provided below:

(1) Using a calendar:
I am a Monday.
I am not a multiple of 4.
My digits sum to 9.
Who am I?

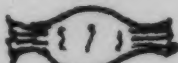
(2) I am a multiple of 2.
I am greater than 20.
I am less than 40.
My digits sum to 8.
Who am I?

Pictorial Problem Types

Graphing. Use graphs that have been collected or teacher created. Make sure to utilize bar, line, circle and picture graphs. Ask pupils questions regarding the pictured results displayed in the graphs.

Pictures to words. This strategy is helpful in moving pupils from pictorial to word-type problem-solving examples. Pupils should be provided with pictorial sequences as shown, and respond first orally and then in writing to the problem and its solution. Pupils also enjoy illustrating and solving their own picture strips.

Pictorial Sequence

Pencil  5¢	Candy  3¢	What's the problem?
---	--	---------------------

Possible responses:

Pencil costs 5¢

Candy costs 3¢

How much spent (in all) = 8¢

How much more does the pencil cost? = 2¢

Word Problem Types

Problems without questions. In this strategy, pupils are provided with written statements and must write their own questions. An example of the problems-without-questions strategy appears below:

Brett and his dad went to the Orioles game.

Tickets cost \$5.00 each. Brett's dad had a \$20. bill.

Possible (written) questions:

Do they have enough money?

How much change will they get?

Who gets the change?

Can they buy popcorn?

Substituting "easier" numbers. Use this strategy by revising a standard word problem to include "workable" quantitative information. After students solve the example with "easier" numbers, have them complete the original example. Consider the example below:

Original Problem--Nita bought 8 pens at 49¢ each, 14 erasers at 19¢ each, and 2 boxes of crayons at \$1.29 each. How much did she pay for her purchases?

Easier Number Problem--Nita bought 3 pens at 50¢ each, 4 erasers at 20¢ each, and 2 boxes of crayons at \$1.00 each. How much did she pay for her purchases?

Problems without numbers. In this strategy pupils should be provided with word examples without numbers. They should then use the basic problem-solving procedure (Figure 3) to assist them in analyzing the problem situation. Consider the following examples:

1. Stacey's mom went to the store to buy some soup. How much did she spend?

What information do you need?
How do you solve this problem?

2. The Yankees lost the game. They played the Orioles. How many more runs did the winning team score?

What information is needed?
Who won?
How do you solve this problem?

Math experience approach. This strategy involves pupils in actually devising and writing problem-solving examples. Begin by having the pupils develop a written problem as a group. Place this problem on a large sheet of chart paper, and solve it as a group. After more experiences in group creation and solution, have the pupils write and solve their own problems.

As indicated in the beginning, this paper is an initial, formative step in devising and developing a model--a sequence and teaching techniques--for problem-solving instruction with low achievers. I hope the ideas presented here will stimulate further interest and research.

References

- Bigge, Morris L. Learning Theories for Mathematics (2nd ed.)
New York: Harper and Row, 1976.
- Fennell, Francis M. Elementary Mathematics Diagnosis and Correction Kit.
Englewood Cliffs, N.J.: Center for Applied Research In Education, 1981.
- Fennell, Francis M. Elementary School Mathematics: Priorities for 1980s.
Bloomington, Indiana: Phi Delta Kappa (fastback) 1981.

The Validation of a Criterion-Referenced Test of Rational Number Concepts and Skills

Barbara R. Sadowski, University of Houston

The back-to-basics movement in mathematics education has stimulated interest in the development of criterion-referenced tests which can reliably discriminate between those students who have mastered a concept or skill from students who have not mastered it. In an earlier report (Sadowski, 1980), an extreme-types latent trait model was described and its use in determining mastery of sets of items was illustrated. Since the extreme-types model is fully described elsewhere (see Dayton and Macready, 1976; Macready and Dayton, 1977) only a brief outline of the assumption of the model and the interpretation of the parameters and test of significance yielded by the data analysis will be described here.

The Extreme-Types Model

The extreme-types latent trait model is based on the assumption that, for an n -item domain for a specific concept, students may be placed in one of two discrete categories: (1) masters of the concept, or (2) nonmasters of the same concept. The extreme-types model attempts to fit the domain scores from a group of S_s to this model while estimating three parameters. The first parameter is θ , the estimated proportion of masters in the domain (i.e., those with a score of n). Two item parameters, α and β are estimated for each item in the domain. α , an estimated intrusion parameter gives the probability that a nonmaster will correctly respond to the item. β , the estimated omission parameter is the probability that a master will not respond correctly to the item. A goodness-of-fit Chi-square statistic is used to determine whether the domain is homogeneous based on the internal consistency of the student responses. The fit of the model to the data is assessed by the Pearson Chi-square goodness-of-fit test. The "expected" or predicted frequencies for each of the score patterns of zeros and ones for the items of a domain are compared to the observed frequencies for each domain. This statistic is evaluated with degrees of freedom equal to $2K - q - 1$, where K = the number of items in the domain, and q is the number of estimated parameters ($q = 9$ for a four-item domain; α and β for each item + θ). Values of p (the right-tail probability of the Chi-square distribution) greater than .001 are evidence of fit to the model.

Construction of the Test Domains

Macready and Merwin (1973) suggest that both item content and internal consistence of student responses be considered when determining the homogeneity of an item domain. For this reason, the all-or-none mastery assumption is more tenable.

The number of items in a domain is determined by (1) the estimated proportion of masters in a population, (2) the reliability of the mastery decision that is desired, and (3) the "guessing rate" of the items. Macready and Dayton show that a mastery score of 3 in a four-item domain will misclassify about 5% of the masters for test items involving little or no guessing when the population is equally divided between masters and nonmasters.

Once the number of items to be constructed for each skill or concept has been determined, the next task is to decide what item content will produce a homogeneous domain when the internal consistency of student responses is considered. For example, although the concept of a fraction as a part of a region should be generalized across circles, rectangles, triangles, and other shapes, research has shown that students recognize parts of circular regions as models for fractions more frequently than other region models. Likewise, the addition of fractions with unlike denominators that are composite numbers have been found to be easier than those with unlike denominators that are prime numbers (Uprichard & Phillips, 1977). Thus, although the test items might be logically assumed to be mathematically homogeneous, the internal consistency of student responses to the items might not support this premise. The test constructor must decide between restructuring the domain or leaving the domain unchanged in order to test a more generalized concept.

The test of rational number concepts and skills used in this study consisted of 21 domains, each having four items. The placement of items into domains was based on an analysis of the item content as determined by the Wilson Content Taxonomy (1976), the skill levels of middle school students, and the methods used in teaching nonnegative rational number concepts and skills. The test consisted of 6 domains (2 addition, 1 subtraction, 1 multiplication, and 2 division). Models for fractions included one domain each on the region model, the set model, the number line model, and the division interpretation of a fraction. Equivalent fraction domains included a region model, a set model, and two domains on renaming to higher and lower terms at the symbolic level only. Comparison of fraction and understanding of terms (improper, proper, and mixed numeral) were each covered by one domain.

Method

The test was administered to all students in the seventh and eighth grades in three middle schools in Houston, Texas (400 Ss). The regular mathematics teacher administered it during the mathematics class period and students were given time to finish all items on the test.

Results

The domains of the region and set models for proper fractions were homogeneous and more than 90% of the 400 students demonstrated mastery of these concepts. In contrast, naming points on a number line as a model for fractions showed almost equal numbers of masters and nonmasters.

The domain of items on the division interpretation of a fraction was not homogeneous, indicating that many students have partial knowledge of the concept. From Table 1, the larger beta value for Item 4 reflects the difficulty students had in naming a fraction ($4 \frac{1}{8}$) on a number line marked off in fourths.

The intrusion parameter (alpha) for Item 2 (locating $\frac{1}{2}$ on the number line) indicates many nonmasters responded correctly. In the division interpretation domain, the larger alpha value for Item 4 indicates that students were able to answer this item (matching $5\overline{)6}$ with $6 \div 5$) while missing Items 1-3 (e.g., matching $5 \div 4$ with $5/4$).

Table 1
Item Parameter Estimates for Number Line Model
and Division Interpretation

Number line model			Division Interpretation	
Theta = .51			Theta = .48	
Item	Alpha	Beta	Alpha	Beta
1	.04	.01	.10	.15
2	.26	.02	.14	.20
3	.03	.13	.19	.25
4	.03	.39	.35	.31

Chi-square = 46.4, $df = 6$ Chi-square = 17.2, $df = 6$

The domains on computation with fractions showed some fairly predictable patterns. Two domains on addition and subtraction are shown here; the estimated parameter values for these domains are in Table 2.

Addition Domain Items

Directions: Compute. Put answers in lowest terms.

(a) $4/5 + 2/3 =$ (b) $1/2 + 3/7 =$ (c) $7/8 + 5/6 =$ (d) $1/2 + 2/3 + 7/8 =$

Subtraction Domain Items

Directions: Compute. Put answers in lowest terms.

(a) $5\ 3/4 - 6/7 =$ (b) $2\ 1/5 - 1\ 3/5 =$ (c) $3 - 1/4 =$ (d) $4 - 2\ 3/8 =$

Table 2

Item Parameter Estimates for Addition and Subtraction

Addition			Subtraction	
Theta = .63			Theta = .64	
Item	Alpha	Beta	Alpha	Beta
1	.00	.16	.05	.26
2	.05	.10	.02	.21
3	.00	.26	.05	.15
4	.00	.31	.08	.18

51

Chi-square = 28.7, $df = 6$ Chi-square = 110.2, $df = 6$

The low alpha values are consistent with Item format while the larger beta values for addition Items 3 and 4 attest to the difficulty of adding more than two fractions or adding fractions with relatively prime denominators. The fact that the correct answer requires writing the answer in lowest terms also affects the difficulty of the problem. This domain is homogeneous, meaning that the masters can be separated from nonmasters with some degree of reliability.

The subtraction domain is not homogeneous, however, and the difficulty of Items 1 and 2 is reflected in the larger beta values. An inspection of the items reveals that these items require different skills from Items 3 and 4. In Table 3 it can be seen that there were 25 Ss who had only Items 3 and 4 correct, and 21 with Items 2, 3, and 4 correct, while 17 were able to do only Items 1 and 2 correctly. Thus, although 101 Ss had scores of 0 and 108 had scores of 4, the domain is not homogeneous.

Table 3
Observed N for Subtraction Domain

Pattern	Observed <u>N</u>	Pattern	Observed <u>N</u>
0000	101	0001	9
1000	6	1001	6
1100	17	1101	5
0010	6	0011	25
0110	5	1011	11
1110	13	0111	21
		1111	108

The two domains on multiplication and division are shown here; the estimated parameter values for these domains are shown in Table 4.

Multiplication Domain Items

Directions: Compute. Put answers in lowest terms.

(a) $1/8 \times 1/2 =$ (b) $2/3 \times 5/6 =$ (c) $3/8 \times 4/5 =$ (d) $3/5 \times 4/7 =$

Division Domain Items

Directions: Compute. Put answers in lowest terms.
 (a) $6/7 \div 5 =$ (b) $3/4 \div 4 =$ (c) $2 \div 3/5 =$ (d) $8 \div 2/3 =$

Table 4
 Item Parameter Estimates for Multiplication and Division

Multiplication			Division	
Theta = .61			Theta = .57	
Item	Alpha	Beta	Alpha	Beta
1	.03	.06	.03	.05
2	.02	.12	.00	.14
3	.02	.14	.01	.26
4	.05	.07	.01	.22

Chi-square = 28.7, df = 6

Chi-square = 88, df = 6

The lack of homogeneity for the multiplication domain was due mainly to a group of 13 Ss who had only Items 1 and 4 correct. An inspection of these two items reveals that the product is found in one step, while Items 2 and 3 require two steps for the correct product in lowest terms. The division domain shows a similar result with Items 3 and 4 requiring one more step in the solution; i.e., changing an improper fraction to a mixed numeral on the last step. Item 3 alone was missed by 21 Ss while 24 Ss were at score 0 and 107 were at score 4. These results are similar to the subtraction domain just cited.

Two domains on comparing fractions are shown here and the estimated parameter values are shown in Table 5.

Comparing Fractions

Directions: Circle the fraction that names the greatest number.
 (a) $1/3$ or $4/8$ (b) $3/4$ or $2/3$ (c) $1/2$ or $2/3$ or $3/4$ (d) $2/5$ or $7/10$ or $1/25$

Least Fractions

Directions: In each set, circle the fraction that names the least number.
 (a) $1/11$, $1/15$, $1/8$ (b) $1/7$, $1/5$, $1/4$ (c) $5/3$, $20/3$, $10/3$ (d) $2/5$, $10/5$, $8/5$

Table 5
 Item Parameter Estimates for Comparing
 Fractions and Least Fractions

Comparing Fractions		Least Fractions	
Theta = .703		Theta = .347	
Alpha	Beta	Alpha	Beta
.351	.071	.421	.00*
.224	.038	.426	.00*
.142	.069	.476	.00*
.213	.164	.490	.00*

Chi-square = 34.8, df = 6

Chi-square = 481, df = 6

*Exceeds 0-1 probability limits

In Table 5, it is apparent, and a not-too-surprising conclusion, that in comparing fractions, two fractions are easier than three. Also, note that the first item, comparing $1/3$ and $4/8$, is much easier than the other three items. There is about 70% mastery on this domain. The domain on choosing the smallest fraction yields data that are uninterpretable because the beta values have exceeded the 0-1 probability limits. From the observed N for patterns shown in Table 6, it is clear why the data do not fit the mastery/nonmastery model.

Table 6
Observed N for Least Fraction Domain

Pattern	Observed <u>N</u> *
0000	59
1100	61
1110	6
1101	12
0011	72
1011	7
0111	7
1111	181

*Observed N for all other patterns is less than 5.

The first two items are apparently different from the last two items since 61 Ss were able to do the first two while 72 Ss could only do the last two. Thus, based on the consistency of student responses, and the fit to the model, the domain is not homogeneous. Partial mastery of the skill of comparing fractions is obvious; a score of 2, however, is not interpretable in terms of the student's mastery skills. Whether the domain should be restructured is a decision left to the test constructor.

Discussion

In applying the extreme-types model to data obtained from a test of fractions taken by middle school students, results are consistent with earlier research. The lack of domain homogeneity across items can be explained by factors such as (1) number of steps in the computation items, (2) the difficulty of supply items compared to multiple-choice items, and (3) the inclusion of "easier" or more familiar fractions such as $1/2$ and $1/4$. Data showed that region and set models for fractions were mastered by almost 100% of the students.

The high mastery percentage on these domains indicates that fewer than four items are necessary to reliably discriminate between the mastery/nonmastery groups. Two items will suffice with 90% mastery in the population. Although not discussed in this paper, domains involving equivalent fractions and renaming improper fractions/mixed numerals had 85% mastery and thus could also be diagnosed with only two items each. Because only 9% of the students had mastered the number line model and division interpretation, the four items are needed, but they need revising to improve discrimination.

Choosing the least (or greatest) of a set of fractions resulted in invalid parameter values reflecting the students' specific knowledge of comparing either unit fractions or fractions with like denominators. Thus, this domain should be separated into two domains having two similar items.

The results of the computation domains were consistent with research showing the number of steps as an important factor affecting the difficulty of the item. Items with mixed numerals were more difficult than the same item with proper fractions. Thus, in order to reliably diagnose different mastery levels in computation with fractions, these factors should be considered in forming domains, particularly in subtraction and division.

Interestingly, the data from the computation domains showed about 60-70% mastery; a proportion that matched the proportion of students who had already reviewed these topics in school. Interviews with the teachers who administered the tests confirmed that most nonmasters were students in classes that had not reviewed (or had never been taught) certain fraction concepts and skills. Thus, the teachers felt the test had done a good job of diagnosing masters from non-masters even for statistically nonhomogeneous domains.

References

- Dayton, C.M., & Macready, G.B. A probabilistic model for validation of behavioral hierarchies. Psychometrika, 1976, 41, 189-204.
- Macready, G.B., & Dayton, C.M. The use of probabilistic models in the assessment of mastery. Journal of Educational Statistics, 1977, 35, 583-598.
- Macready, G.B., & Merwin, J.C. Homogeneity within item forms in domain referenced testing. Educational and Psychological Measurement, 1973, 33, 351-360.
- Sadowski, B.R. Researching the validity of a diagnostic test. In T. Romberg (Ed.), RCDPM 1980 Research Monograph, Kent, Ohio, 115-130.
- Uprichard, A.E., & Phillips, E.R. An intraconcept analysis of rational number addition: A validation study. Journal for Research in Mathematics Education, 1977, 8, 7-16.
- Wilson, J.W. Diagnosis and treatment in arithmetic: Beliefs, Guiding Models and Procedures. College Park: University of Maryland, 1976.

An Exploration: White and Clark's Test of Inclusion Applied to Ex Post Facto Learning Data

Bob Underhill, Virginia Polytechnic Institute
and State University College of Education, Blacksburg, Virginia

Hierarchy validation has been an important research concern for over two decades. From the time of Gagne and Paradise's (1961) early work, considerable discussion and controversy has surrounded this topic.

Several popular statistical techniques were critiqued by White (1974b). He found the following approaches flawed: The Gagne and Paradise index (1961), Walbesser ratios (1968), Resnick and Wang's (1969) use of Guttman's coefficient of reproducibility, and the phi-correlation coefficient (Caple and Jones, 1971). As White (1974b) stated,

All are shown to be unsatisfactory for one or more of the following reasons: sometimes the index can have values which indicate a hierarchical connection even when the skills are really independent; errors of measurement are ignored; there is no way of calculating the probability of error in generalizing from sample to population; and subjective judgments are necessary on how many questions of a set must be correctly answered (p.65).

White and Clark (1973) have developed a test of inclusion which purports to overcome the problems cited with other statistics.

Learning vs. Subskill Retention Hierarchies

Most researchers have examined learning hierarchies. In the sense of Gagne (1962), learning hierarchies represent optimal paths for acquiring concepts and skills. Optimal paths are probably best characterized by their efficiency in time to acquire a superordinate skill. While the time of administration of criterion measures has been a subject of considerable debate, the debate has centered on during vs. immediately following ascension through the specified hierarchy. Gagne and Bassler (1963) noted that many subjects upon successful completion of superordinate tasks no longer retain certain hierarchy subtasks.

The direct application of learning hierarchy statistical tests to ex post facto learning data has been completed by several researchers (Uprichard and Phillips, 1977; Causey, 1976; Hill, 1976). These studies used Walbesser ratios and pattern analysis (Rimoldi and Grib, 1960). Underhill (1978), however, discussed problems related to the use of the "learning hierarchy" label to the data in these studies since typically learners in three or four grade levels were tested over the same concepts. Underhill proposed that the hierarchies be called subskill retention hierarchies since the studies made no attempt to control the time lapse between instruction, reteaching, and data collection. It is quite possible that time ranges differed among subjects by as much as a year.

White and Clark Test of Inclusion

White and Clark (1973) sought to overcome inadequacies of other statistics through the use of the test of inclusion. Three major features of the statistic are (1) errors of measurement are taken into account, (2) probability estimates are calculated, and (3) success criteria are established more objectively than in other statistical tests being used for hierarchy validation.

The test focuses on the number of Ss who answered incorrectly all the questions for the lower skill and correctly all the questions for the higher ones. When this number exceeds a critical value the connection is judged invalid. The critical value is determined by the reliability of the questions and by the value specified for the probability of wrongly rejecting the null hypothesis that the connection is hierarchical (White, 1974a, p. 64).

In the 3 x 3 contingency table presented in Figure 1, the response of every subject must fall into cells 1 through 9 for a two-item response per hierarchy element scheme. Cell 9 is the critical cell. If there are no errors of measurement and if the elements are hierarchically related, cell 9 should be empty. Thus, the test examines whether all respondents of the higher skill are also respondents of the lower skill.

Figure 1

		Number Correct on the Higher Skill			
		0	1	2	
Number correct on the Lower Skill	2	1	2	3	<u>a</u> [= <u>n</u> (1) + <u>n</u> (2) + <u>n</u> (3)]
	1	4	5	6	<u>b</u> [= <u>n</u> (4) + <u>n</u> (5) + <u>n</u> (6)]
	0	7	8	9	<u>c</u> [= <u>n</u> (7) + <u>n</u> (8) + <u>n</u> (9)]
		<u>f</u>	<u>e</u>	<u>d</u>	

The details of the White and Clark Test of Inclusion are presented in the March 1973 issue of Psychometrika. The following comments are intended only to provide a general theoretical framework.

Let $\underline{P}_0$ = the proportion of the population with neither skill

$\underline{P}_1$ = the proportion of the population with only Skill I

$\underline{P}_2$ = the proportion of the population with only Skill II

$\underline{P}_B$ = the proportion of the population with both skills

then $\underline{H}_0: \underline{P}_2 = 0$.

Testing this null hypothesis is complicated by the existence of errors of measurement; thus, estimates must be made of the probability of the number of Ss who could be in $\underline{P}_2$ under $\underline{H}_0$. Since this estimate can only be made when responses are made by Ss to at least two items, two items for each hierarchy element become the a priori minimum for use of this statistic.

The following assumptions are necessary:

- (1) Chance errors on the four questions are independent.
- (2) All scores different from 0 and 2 on each hierarchy level are due to errors of measurement.
- (3) The probabilities of chance errors are equal for the two questions of a given hierarchy element.

Let $\underline{\theta}_a$ = the probability of answering correctly an item on Skill I by a subject who possesses the skill

$\underline{\theta}_b$ = the probability of answering correctly an item on Skill I by a subject who does not possess the skill

$\underline{\theta}_c$ = the probability of answering correctly an item on Skill II by a subject who possesses the skill

$\underline{\theta}_d$ = the probability of answering correctly an item on Skill II by a subject who does not possess the skill

For cell 9, the total probability is the sum of the four theta values multiplied by their respective conditional probabilities, $\underline{P}$.

$$\underline{P}(\text{cell } 9) = \underline{P}_0(1 - \underline{\theta}_b)^2 \underline{\theta}_d^2 + \underline{P}_1(1 - \underline{\theta}_a)^2 \underline{\theta}_d^2 + \underline{P}_2(1 - \underline{\theta}_b)^2 \underline{\theta}_c^2 + \underline{P}_B(1 - \underline{\theta}_a)^2 \underline{\theta}_c^2$$

The θ and P values can be estimated by using marginal totals. White and Clark recommend that the values of $\underline{\theta}_b$ and $\underline{\theta}_c$ take assumed values of 0 and 1, respectively, to make the estimate of $\underline{P}(\text{cell } 9)$ as large as possible; this minimizes the probability of Type I error. Hence, $\underline{H}_0$ can be rejected whenever the observed frequency in cell 9 exceeds a critical value.

Using the assumptions and marginal totals yields the following estimated values:

$$\hat{\theta}_{\underline{a}} = \frac{2\underline{a}}{2\underline{a} + \underline{b}}$$

$$\hat{\theta}_{\underline{c}} = 1$$

$$\hat{\theta}_{\underline{b}} = 0$$

$$\hat{\theta}_{\underline{d}} = \frac{\underline{c}}{\underline{e} + 2\underline{f}}$$

The following equations can be used to calculate other estimated values. $\underline{Q}$ is the estimated proportion of the population with Skill I, and $\underline{R}$ is the estimated proportion of the population with Skill II.

$$\underline{\hat{Q}} = \frac{(2\underline{a} + \underline{b})^2}{4\underline{a}N}$$

$$\underline{\hat{R}} = 1 - \frac{(\underline{e} + 2\underline{f})^2}{4\underline{f}N}$$

$$\underline{\hat{Q}} = \underline{P}_1 + \underline{P}_B$$

$$\underline{\hat{R}} = \underline{P}_2 + \underline{P}_B$$

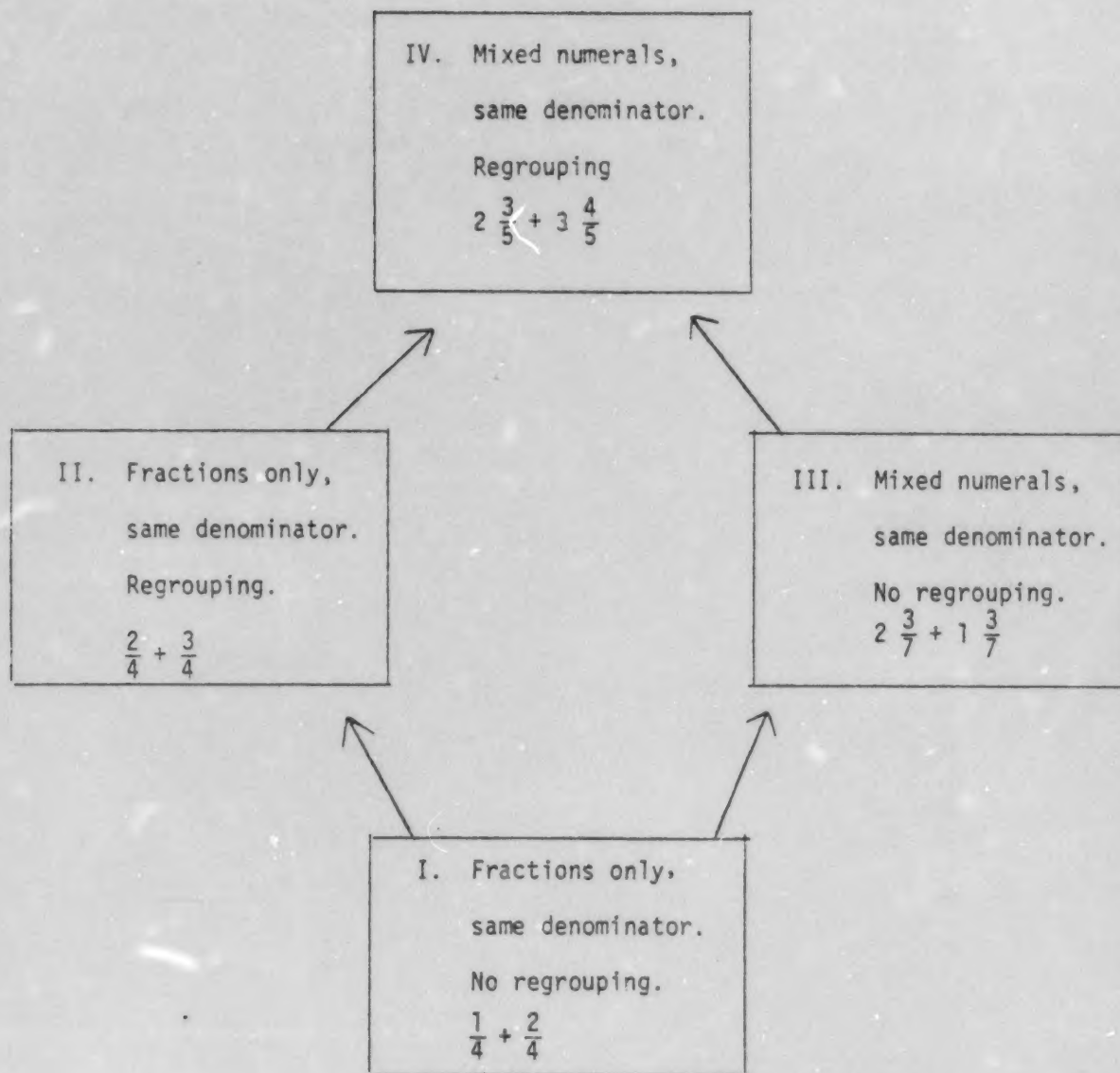
The frequency of cell 9 is a binomial random variable, so the statistical test compares the cumulative probabilities of cell 9 frequencies with a binomial random variable with parameters (N , $\underline{P}$ (cell 9)).

During 1979, data were collected from 325 learners in grades 5-8. The instrument considered of sixteen questions: four each in a four-element hierarchy. The purpose of that investigation was to study the use of Walbesser's criterion ratio on a validated subskill retention hierarchy to determine the problems of applying the ratio to ex post facto learning data and to study the merits of various criterion levels. (Item difficulty was controlled by using only addition combinations of easy and intermediate difficulty, (Murray, 1939)).

The results of the 1979 study indicated that for the hierarchy specified, criterion levels of 2 of 4 and 1 of 2 led to 100% correct conclusions while all other criterion levels (4/4, 3/4, 3/3, 2/3, 2/2, and 1/1) led to one or more incorrect conclusions. Further, the 1/2 and 2/4 criterion levels were affected only slightly by the decrease in consistency ratio values reported by White (1971). A decrease in the values, however, was noted and would, as White predicted, be a major problem in hierarchies that had items with few successful subjects. What was learned, was that the ratio is valid for items with a fairly low level of difficulty and that 1/2 is probably the best success criterion when testing learners of this age to validate subskill retention hierarchies.

Since the bias addressed by White was noted in every success criterion level to some degree, it was concluded that a different statistic might be better suited for the evaluation of subskill retention hierarchies. To this end, the researcher sought to apply the White and Clark test of inclusion to the same data (See Underhill, 1980, for details).

Figure 2



The hierarchy in Figure 2 was used on the basis of task analysis. For each problem type, four examples were constructed. Within each type there was one problem each which used fourths, fifths, sevenths, and ninths. No answers were reducible. From the resulting sixteen items, two successive random orderings were completed to construct each of four test forms. In the first random ordering, each set of four from the first random ordering was randomly ordered to determine the first, second, third, and fourth set of four items on the test sheet. Equal numbers of each of the four forms were used in each classroom of subjects.

Thirteen classes of subjects from grades 5-8 from two elementary schools and one junior high school participated in the study. Based on the performances of these subjects, all subjects with 0 problems correct or sixteen problems correct were excluded from the study as Walbasser (1968) suggested, since they add no data to the question of hierarchical relations. When these subjects were excluded, 140 remained for the analyses.

The responses of every subject who missed all of one level but got at least one correct at a higher level were studied and interpreted to demonstrate that the hierarchical relation was valid in spite of the performances in these few cases. Of the total number of subjects tested ($N = 325$), only 16 such cases were observed. Having done so, the researcher concluded that the four-element hierarchy was valid.

Analysis

The data were subjected to five analyses to examine the validity of decisions of the Index of Inclusion when applied to ex post facto learning data. All analyses used two of the four items for each subskill in the two-item statistical procedure described by White and Clark (1973). Of particular interest were :

(1) the assumption that $\theta_{\underline{b}} = 0$ and $\theta_{\underline{c}} = 1$, and

(2) the criterion levels of $1/2$ for Skill I and $2/2$ for Skill II.

The analyses in Table I were completed.

Results and Conclusions

The results of the five analyses are summarized in Table 2. Four of the analyses for all five hierarchical pairwise orderings confirmed all hierarchical relationships. Only Analysis III failed.

To determine the sensitivity of the test to reject $H_0: P_2 = 0$, the invalid reversed hierarchical pairwise orderings were examined to study under what conditions H_0 would properly be rejected. The results are summarized in Table 3.

Table 1
Analyses Completed

<u>Analysis</u>	<u>Criterion levels</u>		<u>θ assumptions</u>	
	Skill I	Skill II		
I.	1/2	1/2	$\theta_{\underline{b}} = 0$	$\theta_{\underline{c}} = 1$
II.	1/2	2/2	$\theta_{\underline{b}} = 0$	$\theta_{\underline{c}} = 1$
III.	1/2	2/2	None	
IV.	1/2	1/2	$\theta_{\underline{b}} = 0$	$\theta_{\underline{d}} = 0$
V.	1/2	1/2	None	

Table 2

Table of Valid Results from Five Analysis

<u>Skill Order</u>	<u>Analysis I</u>	<u>Analysis II</u>	<u>Analysis III</u>	<u>Analysis IV</u>	<u>Analysis V</u>
1 before 2	+ ^a (11/4) ^b	+ (9/4)	- (4/4)	+ (7/4)	+ (7/4)
1 before 3	+ (3/0)	+ (3/0)	+ (2/0)	+ (0/0)	+ (10/0)
2 before 4	+ (19/4)	+ (13/4)	+ (7/4)	+ (8/4)	+ (18/4)
3 before 4	+ (13/1)	+ (7/1)	+ (4/1)	+ (6/1)	+ (19/1)
1 before 4	+ (7/1)	+ (4/1)	+ (2/1)	+ (4/1)	+ (7/1)

a "+" means "leads to a valid H_0 conclusion"; "-" means "leads to invalid or incorrect H_0 decision."

b Read "The frequency of cell 9 for rejection of H_0 must be 11 or more. Four were observed."

Table 3
Table of Rejections of H_0

	<u>Analysis I</u>	<u>Analysis II</u>	<u>Analysis III</u>	<u>Analysis IV</u>	<u>Analysis V</u>
2 before 1	- (43/29)	- (43-29)	+ (24/29)	- (40/29)	+ (27/29)
3 before 1	- (38/20)	- (30/20)	+ (19/20)	- (31/20)	- (22/20)
4 before 2	- (19/11)	- (25/11)	- (14/11)	- (22/11)	- (23/11)
4 before 3	- (40/14)	- (27/14)	- (16/14)	- (25/14)	- (24/14)
4 before 1	- (43 ^a /42)	- (43/42)	+ (36/42)	- (43/42)	+ (31/42)
2 before 3		- (21/10)	- (11/10)	- (18/10)	- (17/10)
3 before 2		- (18/8)	- (9/8)	- (13/8)	- (16/8)

a Memory overload at $N = 43$. The cumulative probability was 0.21 at termination.

As indicated in Table 3, none of the methods of analysis consistently led to valid decisions. Analyses I, II, and IV were the weakest, leading to erroneous conclusions in every pairwise ordering. Analysis III was the strongest, but even this analysis led to only 3 correct decisions of 7.

Power. White and Clark recommend, to ensure the power is 80% or more, that the observed frequency in cell 9 should be considerably smaller than the expected frequency. Based on the calculations completed by them, a relationship of $f_e > 1.5(f_o + 1)$ was reported as reasonable, i.e., the expected frequency in cell 9 should be greater than 1.5 times the observed frequency plus 1.

Examination of Table 2 leads one to conclude that for a well-defined hierarchy, four of the five analyses will confirm hierarchical relationships when they exist. In the data used in this study, four of the five analyses confirmed all five of the hierarchical pairwise orderings. Only Analysis III failed. The precedence of Skill I before Skill III was incorrectly rejected. Thus, it appears that when one applies the Test of Inclusion to ex post facto learning data that one should, as White and Clark suggested, assume that $\theta_b = 0$ and $\theta_c = 1$ when using 1/2 Skill I and 2/2 Skill II success criteria. This can be readily observed in comparing Analyses II and III. This reduces the probability of rejecting H_0 when it is true. Analysis I, II, IV, and V all appear to minimize the probability of this type of error.

Table 4
Comparison of Analyses

Analysis	Lower Skill	Higher Skill
III	1	2
III	2	4
IV	1	2
V	1	2

Further examination of Table 2 relative to power indicates that the analyses listed in Table 4 do not meet the criterion specified by White and Clark. It appears that analyses III, II, and V lack sufficient power in some instances, even with $N = 140$, to meet the criterion.

Examination of Table 3 leads one to conclude that all five analyses will lead to incorrect acceptance of H_0 when it is false. Analyses III and IV appear to be slightly less susceptible to this problem than analyses I, II, and V.

The scheme in Table 5 summarizes the acceptability (A) or unacceptability (U) based on the previous three paragraphs.

Table 5
Strengths and Weaknesses of Analyses

	Analyses				
	I	II	III	IV	V
Acceptance of H_0 when it is true	A	A	U	A	A
Rejection of H_0 when it is not true	U	U	U	U	U
Estimated power exceeds 80%	A	A	U	U	U

Based on these data, it appears that analyses I and II are best suited for examination of hierarchical relations among ex post facto learning data.

From this study, the following question arises: How does one deal with the probability of accepting H_0 when it is false?

In this study, neither of the two acceptable analyses yielded uniformly valid decisions on this question. There are at least two alternatives. White (1974a) suggests that the decisions on hierarchical orderings be based on logical rather than statistical grounds. This seems to sweep the problem under the carpet. While it is true that logic can be used, it is also true that the frequency of cell 9 can provide some clue: if it is great, immediate reservations are in order. What, however, is "great"? What is a borderline? Further, when dealing with subskill retention, how can one be certain from a diagnostic point of view that the relation is determined by logic? As Uprichard and Phillips (1977) argued in their creation of intraconcept analysis, learning mathematics and remembering mathematics are often a function of psychological and pedagogical variables.

This concern has been addressed by Griffiths and Cornish (1978). They completed a study of three methods for identifying and validating hierarchies. They concluded that the Test of Inclusion provides information for a necessary first step in a two-step validation sequence. They claim that the Dayton and Macready (1976) method of testing hierarchical relationships has two major advantages over the method proposed by White and Clark. First, the Dayton and Macready method considers the hierarchy under examination as one entity, even when the hierarchy is branched. Second, the Dayton and Macready method allows a direct test between hierarchies that contain the same skills in different arrangements. Griffiths, and Cornish, however, suggest that an initial sifting

of the data is often likely to be necessary before the Dayton and Macready method can be applied. This can be accomplished with the test of inclusion.

Further work is needed to extend the work of the present study to other data, age levels, and concepts within the context of diagnostic retention hierarchies. Application of the Dayton and Macready model appears to hold special promise.

Summary

The important goal of evaluating hierarchical dependencies is elusive. The relation between learning tasks and retention tasks was reviewed, and the use of ex post facto data rather than learning data was explored. In the future, researchers interested in validating diagnostic hierarchies can rest assured that the White and Clark Test of Inclusion is among the strongest tests available, and that it appears to be strongest internally when the following are implemented:

- (1) Assume that $\theta_{\underline{b}} = 0$ and $\theta_{\underline{c}} = 1$.
- (2) Establish success criteria of 1/2 on both skills (lower and higher) to yield the best reflections of relationships in the use of the Test of Inclusion when using ex post facto data.

The cell-9 probability, when these assumptions are made, is given by the following steps:

I. Calculate $\hat{\theta}_{\underline{a}}: \hat{\theta}_{\underline{a}} = \frac{2\underline{a}}{2\underline{a} + \underline{b}}$. (See Figure 1 for a and b)

II. Calculate $\hat{\theta}_{\underline{d}}: \hat{\theta}_{\underline{d}} = \frac{\underline{e}}{\underline{e} + 2\underline{f}}$, $\underline{e} = \underline{n}(2) + \underline{n}(5) + \underline{n}(8)$. (See Figure 1)

$$\underline{f} = \underline{n}(1) + \underline{n}(4) + \underline{n}(7). \text{ (See Figure 1)}$$

III. Calculate $\hat{Q}: \hat{Q} = \frac{(2\underline{a} + \underline{b})^2}{4\underline{a}N}$

IV. Calculate $\hat{R}: \hat{R} = 1 - \frac{(\underline{e} + 2\underline{f})^2}{4\underline{f}N}$

V. Under H_0 , $\underline{p}_2 = 0$

VI. Calculate $\hat{\underline{p}}_1$ and $\hat{\underline{p}}_B$: $\hat{\underline{Q}} = \hat{\underline{p}}_1 + \hat{\underline{p}}_B$
 $\hat{\underline{R}} = \hat{\underline{p}}_2 + \hat{\underline{p}}_B$

VII: Calculate $\underline{p}_0$: $\underline{p}_0 = 1 - (\underline{p}_1 + \underline{p}_2 + \underline{p}_B)$

VIII. Calculate the probability of cell-9 being empty:

$$\underline{p}(\text{cell } 9) = \underline{p}_0(\hat{\theta}_{\underline{a}}^2) + \underline{p}_1(1 - \hat{\theta}_{\underline{a}})^2\hat{\theta}_{\underline{d}} + \underline{p}_2 + \underline{p}_B(1 - \hat{\theta}_{\underline{a}})^2$$

- IX. Use a binomial random variable with parameters (N , $P(\text{cell } 9)$) to compare the cumulative probabilities of cell 9 frequencies with the probability of the observed frequency. The following example is paraphrased from White and Clark (1973):

Assume $P_2 = 0.00355$

Assume you get the following values when you use the binomial expansion:

$$\Pr(f_{\text{cell } 9} = 0) = 0.552$$

$$\Pr(f_{\text{cell } 9} = 1) = 0.329$$

$$\Pr(f_{\text{cell } 9} = 2) = 0.097$$

then the cumulative frequency for

$$\Pr(f_{\text{cell } 9} = 1 \text{ or } 2 \text{ or } 3) = 0.552 + 0.329 + 0.097 = 0.978$$

Thus, "an observed value of more than two in (cell 9) would cause rejection of H_0 at the 2 1/2% level" (p. 82).

References

- Capie, W., & Jones, H.W. An Assessment of Hierarchy Validation Techniques. Journal of Research in Science Teaching, 1971, 8, 137-147.
- Causey, M. An Empirically Derived Hierarchy of Intraconcept Relationships: Rational Number Multiplication, (Doctoral dissertation, University of Houston, 1976). University Microfilms No. 77-1503.
- Dayton, C.M., & Macready, G.B. A probabilistic Model for Validations of Behavioral Hierarchies. Psychometrika, 1976, 41(2), 189-204.
- Gagne, R.M. The Acquisition of Knowledge. Psychological Review, 1962, 69, 355-365.
- Gagne, R.M., & Bassler, O.C. Study of Retention of Some Topics of Elementary Non-Metric Geometry. Journal of Educational Psychology, 1963, 54, 123-131.
- Gagne, R.M. & Paradise, M.E. Abilities and Learning Sets in Knowledge Acquisition. Psychological Monographs, 1961, 75, (whole No. 518).
- Griffiths, A.K. & Cornish, A.G. An Analysis of Three Recent Methods for the Identification and Validation of Learning Hierarchies. Toronto, Canada: Paper presented at the Annual Meeting American Educational Research Association, March, 1978.
- Hill, T.A. An Empirically Derived Hierarchy of Intraconcept Relationships: Whole Number Division, (Doctoral dissertation, University of Houston, 1976). University Microfilms No. 77-1513).
- Murray, J. The Relative Difficulty of the Basic Number Facts. In Scottish Council for Research in Education (Ed.), Studies in Arithmetic (Vol. 1), University of London, 1939.
- Resnick, L.B. & Wang, M.C. Approaches to the Validation of Learning Hierarchies. Proceedings of the eighteenth annual regional conference on testing problems. Princeton, J.J., Educational Testing Service, 1969.
- Rimoldi, H.J.A. & Grib, T.F. Pattern Analysis. The British Journal of Statistical Psychology, 1960, 13, 137-149.
- Underhill, R.G. A Diagnosis Position Paper: General Needs and Hierarchies. In J.W. Heddens and F.D. Aquila (Eds.) Proceedings of the Third National Conference on Remedial Mathematics. Kent, Ohio: Kent State University, 1976.
- Underhill, R.G. Some Theoretical and Research Issues Related to Work with Diagnostic Hierarchies. Scottsdale, Arizona: Paper presented at the Fifth National Conference on Diagnostic and Prescriptive Mathematics, April 1978.
- Underhill, R.G. An Empirical Investigation of Walbesser's Consistency Ratio. Vancouver, Canada: Paper presented at Seventh Conference on Diagnostic and Prescriptive Mathematics, April, 1980.
- Uprichard, A.E. & Phillips, E.R. An Intraconcept Analysis of Rational Number Addition: A Validation Study. Journal for Research in Mathematics Education, 1977, 8(1), 7-16.

Walbesser, H.H. An Evaluation Model and Its Application. Washington, D.C.: American Association for the Advancement of Science, 1968.

White, R.T. A Model for Validation of Learning Hierarchies. Journal of Research in Science Teaching, 1974, 11(1), 1-3. (a)

White, R.T. Indexes Used in Testing the Validity of Learning Hierarchies. Journal of Research in Science Teaching, 1974 11(1), 61-66. (b)

White, R.T. & Clark, F.C. A Test of Inclusion Which Allows Errors of Measurement. Psychometrika. 1973, 38, 77-86.

Brain Hemispheric Considerations for a Developmental Approach to Diagnosing Mathematics Concepts for Children

Janice L. Flake, Florida State University

One of the current trends in education is a major emphasis upon diagnosis and prescription. Certainly, one should provide instruction and/or learning experiences that are appropriate for each child. This task is non-trivial. Just exactly what are the key ideas to consider in diagnosing a child, and what are the implications for prescribed instruction? This paper examines a number of background factors, the relationship between stages of cognitive development to brain hemispheric processing, curriculum, some aspects of instruction, and finally to diagnosis and prescription.

One approach to diagnosis has been to use adult logical analysis for breaking down the content into skills continua. On the basis of this analysis, diagnostic instruments have been constructed, testing the carefully written behavioral objectives based on this logical analysis of the curriculum. Some have felt that there exists a linear sequence of behavioral objectives that will span the entire mathematics curriculum, at least at the elementary level. A number of such skills continua currently exist. Why, then, are the children not doing fantastic work in mathematics?

Perhaps other more pertinent questions might be: Is skills mastery the only important criterion to consider? What about the cognitive development of the children? What about the children's attitudes? What about the children's creative abilities? What about the children's thinking processes? And so on ...

Background

Many adults are "turned off" by mathematics. Often they say that they liked mathematics "until we did so-and-so in school", that is until they were exposed to some topic that seemed particularly difficult. Often the enjoyment of mathematics never returned even after the hard topic was past.

Weak areas in mathematics have been identified by several assessment examinations. For instance, the 1972-73 and 1977-78 National Assessments of Educational Progress revealed the following areas of weakness: the concept and use of fractions; problem solving; basic measurement concepts; concepts of perimeter, area, and volume; understanding of geometric topics; estimation, approximation, an aspect of basic "number sense"; and consumer mathematics (Carpenter et al, 1975a, 1975b, 1980a, 1980b).

Brain Research

Another area that seems to be receiving considerable attention in recent times is that of brain research. It is well established that the two parts of the brain (the right hemisphere and the left hemisphere) function in different ways. Some of the key differences are: linear and sequential (left-hemisphere) versus simultaneous and holistic (right-hemisphere) processing; verbal and

explicit (left-hemisphere) versus metamorphic, visual, spatial, and tacit (right-hemisphere) reasoning; and analytical, rational, and logical (left) versus synthetic and intuitive (right) patterns of thinking. Students of mathematics in many classrooms are expected to organize their thoughts in a careful sequential pattern (such as following algorithms or proving theorems) and make explicit definitions. That is, they seem to be presented mainly with left-hemisphere tasks. Is this indeed the case? What about such areas as problem solving, spatial relations, estimating, synthesizing, creativity? Should we, in fact, also be developing these in our mathematics instruction? And if so, are we?

Several (Dake, 1977; Grady and Luecke, 1978; Languis and Kraft, 1977; Olson, 1979; Gowan, 1979; Geiser, 1980; Beals, 1981; Guckes and Elkins, 1981; Torrance, 1981; Zakariya, 1981; and Soares and Soares, 1982) have expressed serious concern that our educational systems have placed the bulk of their emphasis upon left-hemisphere activity, and that very little attention is given to right-hemisphere activities. Thus, we may be educating only half (or perhaps much less) of the brain.

Problem-solving ability has been identified as one of the weak areas in mathematics by National Assessment tests. A positive relationship between problem-solving ability and right hemisphere processing facility has been found (Battista, 1979; Rubenzer, 1979). In addition, positive relationships have been found between problem-solving abilities and : (1) spatial relations (Moses, 1977), (2) instruction in spatial relations (Moses, 1979), and (3) estimation (Hall, 1976). Perhaps this should be considered when mathematics instruction is planned.

Mathematics Achievement

Another area of investigation is the examination of factors that are highly correlated with mathematics achievement. The following factors have been identified:

1. Visual perception (Sawada and Jarman, 1978; Sexton, 1976; Gonyo, 1976)
2. Spatial ability (Guay and McDaniel, 1977, McKay, 1978)
3. Problem-solving ability (Smith, 1978)
4. Conservation abilities (Souveney, 1977; Rohr, 1973; Wheatley, 1967)

(Rohr identified higher levels of conservation such as area and volume to being more closely related to achievement than lower areas such as number.)

5. Field dependence/independence (Anderson, 1976; Wiesner, 1978; McKay, 1978; McLoed and Adams, 1979)
6. Pupil attitudes and self-concept (Pearce, 1976; Hart, 1976)
7. Academic Motivation (Pearce, 1976)

Factors 1 - 3 relate strongly to right-hemisphere activity. Furthermore, Wheatley et al (1978) suggested that conservation tasks of the Piagetian-type (factor 4) perhaps test whether the visual-spatial notions (right-hemisphere activity) have become coordinated with logical thinking (left-hemisphere activity). One way field dependence/independence is tested is by testing one's ability to identify a simple figure embedded in a complex figure. (See Witkin et al, 1962; and Witkin, 1950.) Hence, field dependence/independence (factor 5) measures a coordination of analytic skills and visual perception and space, that is, a coordination between some right-and left-hemisphere processing. Factors 6 and 7 deal with the affective domain, an area Dake (1977) related to right-hemisphere processing.

Thus, each of the factors are connected to right-hemisphere processing. It would seem evident that if we are concerned with mathematics achievement we can not afford to ignore the right hemisphere.

The Young Child

Young children have poor interhemispheric communication (Gazzaniga, 1974; Yacalev and Lacours, 1967). The young child is dominated by right hemisphere orientation (Harris, 1973; and Knox and Kimura, 1970). As language is acquired, the child's thinking pattern slowly shifts to a left-hemisphere orientation (Bogen, 1975). Right-hemisphere cognitive potential is "as important (for high-level problem solving) as language skills" (Bogen, 1975, p. 30), and necessary but not sufficient for creative thinking (Bogen and Bogen, 1969).

Piagetian Stages

Considerable literature now exists describing Piagetian Stages. The following Table will summarize some of their characteristics.

Caution should be used to treat the ages as approximations only, especially with respect to the formal operations stage. A number of studies (Adi, 1976; Arlin, 1975; Cropper, Meck, and Ash, 1977; McKinnon and Renner, 1971; Schwebel, 1975; and Tomlinson-Keasey, 1972) have shown that less than 60% (and considerably less in some cases) of adults have reached the formal operational stage. Perhaps if we had had a curriculum which made full use of right-hemisphere developing and coordination with the left hemisphere the data might be quite different.

The Fifth State--Creativity

A fifth stage of cognitive development has been proposed by Arlin (1975), a stage of creativity. This stage is characterized by achievement of problem finding (versus problem solving) employing the ability to use divergent processes as opposed to convergent processes. The role of the right hemisphere in creativity is discussed in Torrance (1981 and 1982) and Rubenzer (1979).

Table 1

Summary of the Periods of Cognitive Development

Period	Characteristics of the Period	Major Change of the Period
Sensorimotor (0-2 years)	Dominated by sensations and actions, gradual coordination between hand-mouth and hand-eye	Development proceeds from reflex activity to representation and sensorimotor solutions to problems
Preoperational (2-7 years)		
Egocentric stage (2-4 years)	Problems solved through representation-language development (2-4 years); thought and language both egocentric	Development proceeds from sensorimotor representation to prelogical thought and solutions to problems
Intuitive stage (5-7 years)	Cannot solve conservation problems; judgements based on perception rather than logic	
Concrete Operational (7-11 years)	Reversibility attained; can solve conservation problems - logical operations developed and applied to concrete problems; cannot solve complex verbal problems	Development proceeds from prelogical thought to logical solutions to concrete problems
Formal Operations (11-15 years)	Logically solves all types of problems - thinks scientifically; solves complex verbal problems; cognitive structures mature	Development proceeds from logical solutions to concrete problems to logical solutions to all classes of problems

Source: Adapted from B. Wadsworth. Piaget for the Classroom Teacher (New York: Longman, 1978).

Table 2

Stages of Cognitive Development
and
Hemispheric Processing

Stage	Hemispheric Processing
Sensori-Motor	Right hemisphere dominates, only small activity in the left hemisphere
Pre-Operational	Right hemisphere is active; left hemisphere is developing (particularly in the area of language development); poor interhemispheric communication
Concrete Operation	Both hemispheres are active; major emphasis should be coordinating the functions of the two hemispheres (Children should develop a logical understanding of the physical world)
Formal Operation	Emphasis is upon developing the left hemisphere, but should be able to use right hemisphere functions to facilitate left hemisphere activity
Creativity	Dominant activity is in the right hemisphere, but also coordinated with the left hemisphere

Stages of Cognitive Development and Hemispheric Processing

Premature emphasis upon skills continua, computational algorithms, and strictly left-hemisphere activity will ultimately stifle the growth of the human mind. Much greater emphasis should be given to the right-hemispheric activity which appears to proceed as shown in Table 2.

The implications for curriculum are several, and many of them involve the weaknesses identified earlier in this paper. An understanding of fractions (also decimals and percents) builds on an understanding of area and other measurement concepts. In addition, measurement concepts, geometry, estimation/approximation, and problem solving relate strongly to right-hemisphere processing.

It is important, therefore, that children develop concepts such as those listed above by first starting with "hands on" activities of measuring with finding lengths using hand spans and footsteps, covering surfaces, filling containers, comparing weights, estimating measures, and so on.

Examination of Curriculum

The curriculum for early childhood should be rethought. Instead of focusing on reading, writing, and arithmetic, the major thrust of the curriculum of those years should be developing mental imagery and communicating about it. This is important because imagery is the vehicle for connection between the right hemisphere and the left hemisphere, and thus leads to logical thinking.

Concept formation should be clearly distinguished from skill development. The former involves building an understanding of the phenomenon in question, along with developing an imagery about related processes. Skills, on the other hand, are operational in producing answers with speed and accuracy. Instructional strategies for concept formation would be very different from instructional strategies for skill development, with strategies for developing concepts relating more to right-hemisphere type activities, and strategies for skill development would be more left-hemisphere type activities.

Major strands in such a curriculum might include:

1. Classification--how are things alike and how are they different?
2. Seriation--arranging objects from smallest to largest, largest to smallest--first nonverbally, later verbalizing;
3. Patterning--coordinating classification and seriation;
4. Developing number, measurement and spatial concepts, and relationships within and between each of the concepts of number, measurement and space.

Suggestions for Instruction

Processes of instruction with emphasis upon concept formation have been considered by many educational leaders (e.g., Bruner, 1956, 1966, 1971; Dienes, 1971; Skemp, 1971; Beberman, 1971). Two areas of instruction--uses of haptic experiences and uses of computers--offer special promise.

Haptic Experiences

Increasing attention is being given to the role of haptic experience, that is, experience of recognizing objects through active manipulation in the absence of visual stimulation. While we are predominantly a sight-dominated society, a number of people (Lowenfeld, 1945 and 1957; Dorethy, 1975; Piaget and Inhelder, 1956; Piaget et al, 1970; Derevensky, 1978; Wolff, 1972; Guarniero, 1977; and Butzow and Schlenker, 1978) feel that spatial relations and visual perception are at least partially gained through haptic experiences.

What is the role of tactual and visual experiences? Those having both tactual and visual experiences gained considerably more than those having only one, tactual or visual, according to Wolff (1972). The author of this paper has conducted several experiments in which the removal of visual stimulation forced subjects to use other senses. Children and adults who took part in blindfold experiences also wanted to "sneak a peak" after tactually exploring materials. When they did this, they felt they "saw" far more than if they had just the visual experience alone.

Developmental levels for the ability to become better searchers and exploration for forming more accurate mental images to be formed has been investigated (Peel, 1959; Laurendeau and Pinard, 1970; Fisher, 1962; Page, 1959; and Piaget and Inhelder, 1956). Evidence revealed that perceptual activity increased as the child's developmental level increased.

The right hemisphere plays an important role in haptic experiences (Harris, 1975; Milner and Taylor, 1972; Witelson, 1974; Butzow and Schlenker, 1978; and Johnson, 1982). Further, the use of the left hand in exploring materials was found to be superior in exploration than the right hand for all ages (Derevensky, 1978). Hermelin and O'Connor (1971a and 1971b) found that blind children and adults tend to rely more on their left hand in a Braille reading task.

Research would suggest that as students explore physical materials, they should:

1. Learn to discriminate various features;
2. Recognize that which they are exploring;
3. Be able to reproduce in great detail that which they are exploring, thus indicating that a detailed image has been formulated.

Teachers who seek to establish mental images of concepts, then should plan instructional sequences which include haptic exploration.

Uses of Computers

Within the past few years, computers have become very popular. While, there likely are more materials available in elementary mathematics than in any other educational area, most of these materials ignore right-hemisphere development. Torrance (1981) has expressed concern that computer-based materials are tied to linear-type programmed learning materials, and suggested that greater emphasis should be devoted to non-linear programs that can increase the development of the right hemisphere.

Emphasis should be given to uses of the computer for right-hemisphere development in both uses of prepared materials and in developing programming abilities. A program designed for right-hemisphere development (Flake, 1983b) has the student transform various shapes on the video screen using the keyboard and game paddles by sliding, turning, flipping, and/or dialing it.

Papert (1980) has suggested very inventive ways of using the computer through the use of LOGO, a language designed for early entry into computer programming. Even young children can use it to make graphic displays by giving the computer commands. Such experiences will encourage the child to integrate right-and left-hemisphere activities since the child must break down the mental picture that he/she wishes to construct into component parts in order to write a program for each part. The left-hemisphere activity is therefore dependent on the right-hemisphere activity.

Logical inferences, problem solving, formal operations, visual ratio and proportion, and spatial-visualization of people who had spent a year in computer programming appear to be dependent upon spatial-visualization ability (Flake, 1982 and 1983a). Further, a hierarchical sequence (from lowest to highest) of spatial-visualization, visual ratio and proportion, and understanding computer graphics appears to exist.

Implications for Diagnosing Children's Mathematical Concepts

The experiences of children in the elementary school affect their overall development. Children should have considerable right-hemisphere activity and eventually a coordinating between right and left hemispheres. The shift to left hemispheric activity should be gradual, rather than abrupt--as is usually done in many classrooms. More emphasis should be given to conceptual development during these years rather than primary emphasis upon skills.

Diagnosis should be more focused on major stages. Prescriptions should identify more classes of activities/experiences for which the children can develop more simultaneous (non-linear) and holistic type thinking and gradually bring organization to it, rather than linear skills continua.

Effective diagnosis can prevent problems. It need not be concerned with remediation only. The major question should be: When has the child conceived an image of models representing the concept in question? In order to test this, logical questions about concrete objects should be asked before and after they are physically changed. Children unable to consider logical questions about the objects should continue to work with the objects rather than moving on to more abstract forms of instruction.

Piaget identified the following factors for intellectual development: (1) maturation, (2) experience, (3) social transmission, and (4) self-regulation. As teachers, we can do nothing about maturation except recognize that it exists. Experience, however, is another matter; it should start with right-hemisphere activities. Social transmission means allowing children to interact with each other and the teacher. Many diagnostic/prescriptive programs stress individualized instruction. It is crucial, however, to distinguish between "individual" and "individualized" instruction. The fourth factor, self-regulation, can evolve out of holistic approaches as opposed to skills continua.

Factors to consider in diagnosing and prescribing instruction in mathematics for children should include a number of aspects, such as hemispheric research, the possible relationship between stages of cognitive development and hemispheric processing, curriculum and instruction. The role of the right hemisphere appears to be more important than is often treated in many school programs, since learning seems to start in the right hemisphere.

References

- Adi, H. The interaction between the intellectual development levels of college students and their performance on equation solving when difference reversible processes are applied. Ph.D. dissertation. Florida State University, 1976.
- Anderson, W.L., Jr. A study to determine the correlation between academic achievement and field dependence-independence in third, fourth, fifth, and sixth grade pupils. Ed.D. dissertation. Brigham Young University, 1976.
- Arlin, P. Cognitive development in adulthood: a fifth stage? Developmental Psychology, 1975, 11, 602-606.
- Battista, M. Interrelationships between problem solving ability, right hemisphere processing facility, and success in mathematics instruction. Paper presented at the fifty-seventh annual conference of the National Council of Teachers of Mathematics. Boston, MA, 1979.
- Beals, M.G. The right brain: an emerging frontier in education. Paper presented at the Annual Symposium "Education and Contemporary America" of the Boise State University School of Education, October 1981.
- Beberman, M. UICSM Looks at elementary mathematics. Mathematics Teaching, 1971, 55, 26-28.
- Bogen, J.E. Educational aspects of hemispheric specialization. UCLA Education, 1975, 17, 23-33.
- Bogen, J.E. and G. M. Bogen. The other side of the brain III. The Corpus Collosum and Creativity. Bulletin of the Los Angeles Neurological Society, 1969, 34, 191-220.
- Bruner, J.S.; J.J. Goodnow; and G.A. Austin. A Study of Thinking. New York: John Wiley & Sons, 1956.
- Bruner, J.S. On Knowing. New York: Atheneum (Originally published by Harvard Press), 1971.
- Bruner, J.S. Toward a Theory of Instruction. New York: W.W. Norton & Company, 1966.
- Butzow, J.W. and R.M. Schlenker. A study of a possible relationship between Lowenfeld's visual-haptic theory and Piaget's developmental theory. Paper presented at the National Science Teaching Convention. Toronto, Ontario, March, 1978.
- Carpenter, T.P.; Coburn, T.G.; Reys, R.E.; Wilson, J.W. Results and implications of the NAEP mathematics assessment: elementary school. The Arithmetic Teacher, October 1975a, 22:6, 438-450.
- Carpenter, T.P.; Coburn, T.G.; Reys, R.E.; Wilson, J.W. Results and implications of the NAEP mathematics assessment: secondary schools. The Mathematics Teacher, October 1975b, 68:6, 453-470.

- Carpenter, T.P.; Kepner, H.; Corbitt, M.K.; Linguist, M.M.; and Reys, R.E. Results and implications of the second NAEP mathematics assessments: elementary school. The Arithmetic Teacher, April 1980a, 27:8, 10-12, 44-47.
- Carpenter, T.P.; Kepner, H.; Corbitt, M.K.; Linguist, M.M. and Reys, R.E. Results of the second NAEP mathematics assessment: secondary school. The Mathematics Teacher, May 1980b, 73:5, 329-338.
- Cropper, D.A.; Meck, D.S.; and Ash, M.J. The relation between formal operations and a possible fifth stage of cognitive development. Developmental Psychology, 1977, 5, 517-518.
- Dake, M. An education strategy for the right hemisphere of the brain. Paper presented at the Ninth Annual Meeting of the International Visual Literacy Association. Iowa City, Iowa, May 1977. ERIC ED 154 432.
- Deverensky, J.L. A developmental study of haptic perception in young children. Paper presented at the Annual Meeting of the American Educational Research Association. Toronto, Canada, March 1978. ERIC ED 152 407.
- Dienes, Z.P. Building Up Mathematics. London: Hutchinson Educational, 1971.
- Dorethy, R.E. Art education and perceptual theory: a critical analysis, 1975. ERIC ED 147 266.
- Fisher, G.H. Objects, models, and two-dimensional shapes for use in visual and tactile-kinaesthetic shape perception experiments. Research Report No. 1. Institute of Education. University of Durham, 1962.
- Flake, J.L. Report to Broward County on Uses of Microcomputers Part II: Technical Report. Report prepared as part of a Florida State University/Broward County Teacher Education Center contract, 1982.
- Flake, J.L. Some intellectual processes involved in computer programming. Publication in progress, 1983a.
- Flake, J.L. Transformational geometry and the microcomputer. Publication in progress, 1983b.
- Gazzaniga, M.S. Cerebral dominance viewed as a decision system. In S.J. Dimond and J.G. Beaumont (eds.). Hemisphere Function in the Human Brain. New York: Wiley and Sons, 1974.
- Geiser, W.R. Authority without wisdom. Community College Frontiers, Spring 1980, 8:3, 49-52.
- Gonyo, M.E. The relationship between visual perception and arithmetical computation skills among learning disabled second and third grade children. Ed.D dissertation. Rutgers University The State University, 1976.
- Gowan, J.C. Are we educating only half of the brain? Gifted Child Quarterly, Spring, 1979, 23:1, 6-15.
- Grady, M.P. and E.A. Lueke. Education and the brain. Bloomington, Ind.: Phi Delta Kappa Education Foundation, 1978. ERIC ED 153 258.

- Guarniero, G. The senses and the perception of space. Ph.D. dissertation. New York University, 1977.
- Guay, B.G. and E.D. McDaniel. The relationship between mathematics achievement and spatial abilities among elementary school children. Journal for Research in Mathematics Education, May 1977, 8:3, 211-215.
- Guckes, L. and R. Elkins. Implications of brain research for educational practices. A paper presented at the Annual Symposium "Education and Contemporary America", of the Boise State University School, October 1981.
- Hall, W.D. A study of the relationship between estimation and mathematical problem solving among grade students. Ed.D. dissertation. University of Illinois, 1976.
- Harris, L.J. Neurophysiological factors in the development of spatial skills. In J. Eliot and N.J. Salkind (eds.): Children's Spatial Development. Illinois: Charles C. Thomas, 1975.
- Harris, L.J. Neurophysiological factors in spatial development, 1973. ERIC ED 036 163.
- Hart, K.M. Mathematics achievement and attitudes of nine and ten-year olds, effects of mathematical games and puzzles. Ed.D. dissertation. Indiana University, 1976.
- Hermelin, B. and N.O'Connor. Functional assymetry in the reading of Braille. Neuropsychologia, 1971a, 9, 431-435.
- Hermelin, B. and N.O'Conner. Right and left handed reading of Braille. Nature, 1971b, 231, 470.
- Johnson, V.R. Myelin and maturation: a fresh look at Piaget. Science Teacher, March 1982, 49:3, 41-44, 49.
- Knox, C. and D. Kirmura. Cerebral processing of verbal and non-verbal sounds. Neuropsychologia, 1970, 8, 22.
- Languis, M. and R.H. Kraft. An educational perspective on the hemisphere process of the brain. Paper presented at the Twenty-second Annual Meeting of the International Reading Association. Miami Beach, May 1977. ERIC ED 151 748.
- Laurendeau, M. and A. Pinard. The Development of the Concept of Space in the Child. New York: International Universities Press, 1970.
- Lowenfeld, V. Creative and Mental Growth, 3rd ed. New York: MacMillan, 1957.
- Lowenfeld, V. Tests and visual and haptic aptitudes. American Journal of Psychology, 1945, 58, 100-111.
- McKay, L.R. Sex differences in cognitive styles and mathematics in fourth and eighth graders. Ph.D. dissertation. University of Texas at Austin, 1978.

- McKinnon, J. and J. Renner. Are colleges concerned with intellectual development? American Journal of Physics, September 1971, 39, 1047-1052.
- McLeod, D.B. and V.M. Adams. The interaction of field independence with small group instruction in mathematics. Paper presented at the Fifty-Seventh Annual Conference of the National Council of Teachers of Mathematics. Boston, MA. 1979.
- Milner, B. and L. Taylor. Right-hemisphere superiority in tactile pattern-recognition after cerebral commissurotomy: evidence for nonverbal memory. Neuropsychologia, 1972, 10, 1-15.
- Moses, B. The effects of spatial instruction on problem-solving performance. Paper presented at the Fifty-Seventh Annual Conference of the National Council of Teachers of Mathematics. Boston, MA, 1979.
- Moses, B. The nature of spatial ability and its relationship to mathematical problem solving. Ph.D. dissertation. Indiana University, 1977.
- Olson, M.B. Cerebral lateralization in science. The Gifted Child Quarterly, Spring 1979, 23:1, 142-150.
- Page, E.I. Haptic perception: a consideration of one of the investigations of Piaget and Inhelder. Education Review, 1959, 11, 115-124.
- Papert, Seymour. Mindstorms: Children, Computers, and Powerful Ideas. New York: Basic Books, 1980.
- Pearce, W.C., III. The characteristics of the low-achieving mathematics student entering selected community college in East Texas. Ed.D. dissertation. East Texas State University, 1976.
- Peel, E.A. Experimental examination of some of Piaget's schemata concerning children's perception and thinking, and a discussion of their educational significance. British Journal of Educational Psychology, 1959, 29, 89-103.
- Piaget, J.; B. Inhelder; and A. Szeminska. The Child's Conception of Geometry. New York: Basic Books, Inc., 1970.
- Piaget, J. and B. Inhelder. The Child's Conception of Space. London: Rutledge, Kegan Paul, 1956 (originally published, 1948).
- Rohr, J.A.G. The relationship of the ability to conserve on Piagetian Tasks to achievement in mathematics. Ed.D. dissertation, The University of Tennessee, 1973.
- Rubenzer, R. The role of the right hemisphere in learning & creativity implications for enhancing problem solving ability. The Gifted Child Quarterly, Spring 1979, 23:1, 78-100.
- Sawada, D. and R.F. Jarman. Information matching within and between auditory and visual sense modalities and mathematics achievement. Journal for Research in Mathematics Education, March 1978, 9:2, 126-136.

- Schwebel, M. Formal operations in college freshman. Journal of Psychology, 1975, 91, 133-141.
- Sexton, L.C. Auditory and visual perception, sex, and academic aptitude as predictors of achievement for first grade children. Ed.D. dissertation. Ball State University, 1976.
- Skemp, R.R. The Psychology of Learning Mathematics. Middlesex, England: Penguin Books Ltd., 1971.
- Smith, R.L. Figure types and problem-solving achievement in computational activities. Ph.D. dissertation. The Ohio State University, 1978.
- Soares, L.M. and A.T. Soares. Brain research: Implications for learning. Paper presented at the Congress of Applied Psychology. Edinburgh, Scotland, July 1982.
- Souviney, R.J. The relationships between cognitive operation competencies and computational achievement. Ph.D. dissertation. Arizona State University, 1977.
- Tomlinson-Keasey, C. Formal operations in females for 11 to 54 years of age. Developmental Psychology, 1972, 6:2, 364.
- Torrance, E.P. Implications of whole-brained theories of learning and thinking for computer-based instruction. Journal of Computer-Based Instruction, May 1981, 7:4, 99-105.
- Torrance, E.P. Hemisphericity and creative functioning. Journal of Research and Development in Education, Spring 1982, 15:3, 29-37.
- Wadsworth, B.J. Piaget for the Classroom Teacher. New York: Longman, 1978.
- Wheatley, G.H., Jr. Conservation, counting, cardinality as factors in mathematics achievement among first grade students. University Microfilms (Order No. 68-15556), 1967.
- Wheatley, G.H., Jr.; R. Franklin; R. Mitchell; and R. Kraft. Hemispheric specialization and cognitive development: implications for mathematics education. Journal for Research in Mathematics Education, 1978, 9, 20-32.
- Wiesner, D.S. Cognitive styles of information processing as related to academic achievement of third grade pupils. Ed.D. dissertation. The University of Rochester, 1978.
- Witelson, S.F. Hemispheric specialization for linguistic and non-linguistic tactual perception using a dichotomous stimulation technique. Cortex, 1974, 10, 3-17.
- Witkin, H.A. Individual differences in ease of perception of embedded figures. Journal of Personality, 1950, 19, 1-15.
- Witkin, H.A.; Dye, R.B.; Faterson, H.F.; Goodenough, D.R.; Karp, S.A. Psychological Differentiation. Studies of Development. New York: Wiley and Sons, 1962.

Wolff, P. Action and imagery in children. In Levin, J.R.; Wolff, P.; Davidson, R.E.; and Rohwer, W.D., Jr. Issues in Imagery Learning: Four Papers. Theoretical Paper No. 36. Wisconsin Research and Development Center for Cognitive Learning, 1972.

Yacalev, P. and A.R. Lacours. The myelogenetic cycles of regional development of the brain. In A. Minowski (ed.). Regional Development of the Brain in Early Life: Symposium. Philadelphia: F.A. Davis, 1967.

Zakariya, S.B. His brain, her brain: a conversation with Richard M. Restak. Principal, May 1981, 60:5, 46-52.

Doing It Their Way: Some Child-Methods in Mathematics

Lesley R. Booth and Kathleen Hart, Chelsea College, University of London

Two projects funded by the Social Science Research Council and based at Chelsea College have been concerned with investigating children's understanding of different mathematical topics, and with identifying and analysing some of the errors which many children make. The Concepts in Secondary Mathematics and Science (CSMS) project, which spanned five years from 1974 to 1979, set out to determine how easy or difficult given kinds of mathematical concepts were for children aged 11 to 16 years. Describing a hierarchy of concepts or "levels of understanding" in different topic areas of mathematics, such as fractions, ratio, graphs, algebra and measurement was a primary goal of the study.

Besides describing such a hierarchy in each of the topic areas studied, the CSMS project also identified certain errors which large numbers of children were making. Consequently a second research project, the Strategies and Errors in Secondary Mathematics (SESM) project investigated the reasons for these errors in an attempt to answer the question "What are children actually doing when they produce these wrong answers?" This question is of interest not only because it might help determine ways of assisting children to avoid making these particular kinds of error, but also because it might provide insights into why some children advance slowly in their understanding of mathematics in general. The areas chosen for study included ratio, algebra (generalised arithmetic), fractions, graphs and measurement. In order to investigate the causes of the errors, individual interviews were conducted with children who had been shown to be making the identified errors. During the interviews the children were asked to say what they thought the questions meant, and how they would work out and write down the answers.

From these interviews it became apparent that children often develop methods for solving problems which are less formalised than those typically taught in the mathematics classroom (cf. Erlwanger, 1975; Ginsburg 1975). These "child-methods" (Booth, 1981a; Hart 1981b) which enable the child to successfully solve simple examples, are common sense methods. They are often based upon first principles, largely dependent upon the particular context of the question, rarely symbolised by the child, and based mainly upon the operations of counting, adding and combining with whole numbers and "halves". While appropriate to the solution of easy problems, however, these methods are of limited applicability and do not readily extend to more complex questions or questions in which the numbers are large or non-integer. These child-methods may both bring the child success on easy items and cause the child to make particular errors on items to which the methods may be inappropriate.

Types of child-methods, their usefulness to the solution of easy problems, and the way in which they may prove ineffectual when the problem becomes harder, can be illustrated by instances which have received particular attention in connection with the SESM work on generalised arithmetic. Some children, for example, have difficulty in selecting the correct arithmetical expression to match a given problem, as is required by the CSMS Number Operations test item shown in Figure 1. In some cases, of course, the problem relates to how the division should be recorded. But in other cases the difficulty lies in recognising that the problem requires a division operation in the first place. If

asked to solve the problem, however, these children are often able to do so, especially if the numbers involved are small. The process of solution, however, is not the formal mathematical process of considering the structural nature of the problem and selecting the appropriate mathematical model, but rather a more informal "sharing out" process:



A gardener has 391 daffodils. These are to be planted in 23 flowerbeds. Each flowerbed is to have the same number of daffodils. How do you work out how many daffodils will be planted in each flowerbed?

$391 - 23$	$23 \div 391$
$23 - 391$	391×23
$391 + 23$	$23 + 23$
23×17	$391 \div 23$

Figure 1

- YG (12 yrs): You er ... I know what to do but I can't say it ...
- MB: Yes, well you do it then. Can you do it?
- YG: Those are daffodils and these are flowerbeds, large you see... Oh! they're being planted in different flowerbeds, you'd have to put them in groups ...
- MB: Yes, how many would you have in each group? What would you do with 23 and 391, if you had to find out?
- YG: See if I had them, I'd count them up ... say I had 20 of each... I'd put 20 in that one, 20 in that one ...
- MB: Suppose you had some left over at the end when you've got to 23 flowerbeds?
- YG: I'd plant them in a pot (!) (Brown & Kuchemann, 1976, p.15).

This latter method, while appropriate to a large range of questions, becomes cumbersome when the numbers are larger, and does not generalise to division problems where the sharing-out model may be inappropriate.

Similarly, when asked to record the area of a rectangle, some children do not approach the problem formally (by means of the rule or formula), but rather count the "squares" contained within the figure instead. In this instance the child-method used relates to the way in which the concept of area measurement is first introduced to the child, and almost certainly reflects a perseverance of that early notion. The method is again one which can be successfully applied in the case of simple examples (Figure 2a), but which may lead to mistakes when the numbers are larger and there are consequently many squares to count (Figure 2b), and which is totally inappropriate to the generalised example (Figure 2c).

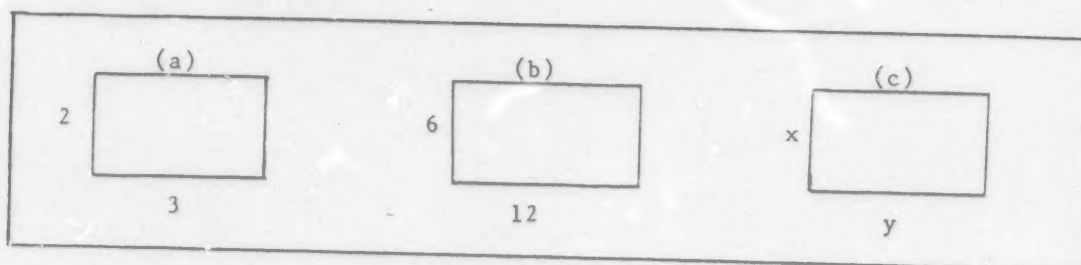


Figure 2

The child who has only the square-counting method available can cope with the generalised problem (2c) only when numerical values are ascribed to the letters involved, a technique which is often adopted by children confronted with algebraic problems of this kind (Kuchemann, 1978; Kuchemann in Hart, 1981a).

Many of the difficulties which children have with "harder" mathematical problems may relate to the child's use of non-formalised and informal methods. In particular, such non-formalised and non-symbolised methods may be expected to lead to difficulty in algebra, where the representation of formal problem structure is often required. Children who do not have the required mathematical structure available in the arithmetic case will find it difficult to produce (or even understand) the generalised form of that structure in algebra.

The results of the individual interviews which SESM has conducted in "generalised arithmetic" or beginning algebra certainly seem to support this view (Booth, 1981b). At the same time, other difficulties relating to the symbolising of method have also become apparent. For example, children may in fact use an appropriate method in the arithmetic case, but symbolise or record it in a manner which is not appropriate to algebraic representation:

- Q: West Ham scores x goals and Manchester United scores y goals. What can you write for the number of goals scored altogether?
- WM (14 years): (Pause) Don't know. It's the x and y , I don't know how many goals they are.
- Interviewer: Right, If you did know how many goals they are, how would you find the number of goals scored altogether?
- WM: By adding them up.
- Interviewer: So you can add those?
- WM: No, because they're not numbers.
- Interviewer: Could you write down what you'd do?
- WM: What, adding them up? Well, what am I adding up!
- Interviewer: Suppose West Ham scored 2, and Manchester United scored 5 ...
- WM: 7 goals. (Writes "7 goals")
- Interviewer: And can you write down what you do to get 7?
- WM: (Laughs) Just add them up! You don't write it down!
- Interviewer: Okay, let's say West Ham scored 73 goals and Manchester United scored 96 ...
- WM: Well, you add them (writes) 73

$$\begin{array}{r} 73 \\ +96 \\ \hline 169 \end{array}$$

Interviewer: And can you write anything down with the x and y, to show you're adding them?
 WM: No! (Laughs)

Another student was able to write something down at this stage:

JD (14 years): Well, I suppose you'd just do x plus y. (Writes $x + y$)
 (Pause. Then as an afterthought writes in "z" as the answer.)
 Interviewer: What have you done now?
 JD: Well, I don't know what the answer is. It's another number, so I suppose you just put a different letter. If you add two numbers you get another number so I suppose adding two letters gives you another letter.

This example nicely demonstrates JD's understanding of the concept of closure in mathematical systems, but unfortunately is not usually regarded as correct in algebra!

Children may both use an appropriate formal method in arithmetic and record it in a manner suitable to algebraic representation yet still have difficulty with algebraic problems of the above kind. In this case the misunderstanding may be of a more general type. For many children, mathematics is an empirical subject which relates to the derivation of specific numerical answers. For such children, a "question" is distinct from an "answer", and the representation of what (to the child) is a "method" may not be regarded by the child as appropriate to the solution of algebraic problems. Consequently the child will write nothing at all, assuming that something else must have been intended, or will write the generalised statement and then attempt to derive an answer from it. In either case, the child's proper understanding of the problem itself is masked by a misunderstanding of what might legitimately constitute an answer.

The main conclusions derivable from these studies are that children may have ways of solving problems other than those methods taught in the classroom and that children may have difficulty in representing or symbolising procedures even when these are the formal "taught" ones. If it is desired that children learn the formal methods, teachers must first recognise that children may have devised alternative procedures, and then help children to reflect upon them, assessing their uses and their limitations. Perhaps only when children are aware that their own common-sense methods do have limitations will they be persuaded to attend to the more formal taught procedures. At this stage it will also be necessary to demonstrate the relationship between the child's approach and the formal method, so that the latter will acquire a firmer conceptual basis. The difficulties children have with formal symbolic representation of mathematical procedures must not be underrated. If children do not fully understand how expressions in arithmetic should be symbolised, they may have little chance of comprehending algebraic representation, or the purpose for it.

These, and other areas of difficulty revealed by the interviews, are being taken as the basis upon which teaching modules are being developed by SESM. By addressing the problems which appear to give rise to the identified errors, it is hoped that the modules will help children to acquire a better understanding of the topic in question, and hence to avoid making the errors.

References

- Booth, L.R. Child-Methods in secondary mathematics. Educational Studies in Mathematics, 1981, 12, 29-40. (a)
- Booth, L.R. Strategies and errors in generalised arithmetic. Proceedings of the Fifth Conference of the International Group for the Psychology of Mathematics Education, Grenoble, 1981, 140-146. (b)
- Brown, M. and Kuchemann, D.E. "Is it an 'add', Miss?" Part I. Mathematics in School, 1976, 5 (5), 15-17.
- Erlwanger, S.H. Case Studies of children's conceptions of mathematics--Part I. Journal of Children's Mathematical Behaviour, 1975, 1 (3), 157-283.
- Ginsburg, H. Young children's informal knowledge of Mathematics. Journal of Children's Mathematical Behaviour, 1975, 1, 63-156.
- Hart, K. (Ed.) Children's understanding of mathematics: 11-16. London: Murray, 1981. (a)
- Hart, K. Investigating understanding. Times Educational Supplement, March 27, 1981, 45-46. (b)
- Kuchemann, D.E. Children's understanding of numerical variables. Mathematics in School, 1978, 7 (4), 23-26.

Diagnostic and Prescriptive Instruction Applied to Teaching Mathematically Talented Students

Kevin G. Bartkovich, Talent Identification Program, Duke University

Introduction

Diagnostic testing and prescriptive instruction in mathematics is usually assumed to involve students who have difficulty learning specific topics. Presently, this assumption is especially prevalent because of the recent emphasis on proficiency in basic skills and the resulting attempts to raise the mathematics knowledge of all students to a minimal competency level. Diagnostic testing is certainly a very useful ingredient in remediation efforts. It is also an extremely effective tool when applied to the other end of the learning spectrum, namely, to those students who excel in mathematics. The use of diagnostic testing followed by prescriptive instruction (DT→PI) with the highly talented mathematics student can significantly increase the amount of material covered and allow for concepts to be studied in greater depth.

Philosophy of Instruction

The multitude of applications of mathematics to other disciplines makes mathematics an especially valuable subject in our schools. Not only do the sciences and engineering have a crucial dependence on mathematics, but also the behavioral and social sciences. The latter have integrated many mathematical concepts into their research, especially through quantitative psychology, economics, statistics, and computers. The close relationship of mathematics to many disciplines implies that the educational opportunities extended to highly able mathematics students will have far reaching consequences in the future. The quality of mathematics education provided for these youths not only will affect the caliber of future mathematicians, but also the ability level of our future leaders in science, engineering, medicine, and other quantitative disciplines. The development of mathematically gifted youths is crucially important in a technologically advanced and dependent society such as ours.

What should be done to maximize the potential of the mathematically gifted? Conant (1959, p.59), former president of Harvard University, addressed the needs of academically talented students in his 1958 study entitled The American High School Today. Concerning this group he stated:

To my mind the most compelling argument is that the student in question has potentialities shared with only a relatively few contemporaries, probably not more than 15 percent of his age group. If these potentialities are not developed as far as possible during the school years, they may never be fully developed. From the point of view of the individual, failure to develop talent in high school may be the equivalent of locking many doors. For example, without mathematics and science in high school, it would be difficult later to enter an engineering school, to take a premedical course in college, and impossible to begin a scientific career in a university.

In addition to the upper 15 percent whom he considered to be academically talented, Conant went on to designate the upper three percent in ability nationwide as "highly gifted pupils." He recommended that identification of these youths take place in the seventh or eighth grade so that an especially challenging secondary school program could be designed. He asserted that the mathematics program for the highly gifted pupils be such that they take Advanced Placement Calculus during twelfth grade (Conant, 1959, pp. 62-63).

Stanley, of The Johns Hopkins University, has found through the Study of Mathematically Precocious Youth (SMPY) that youths in the upper one percent in mathematical reasoning ability are capable of even greater achievements than Conant suggests. There is much documentation (Stanley, 1976) showing that such students are able to complete the precalculus mathematics sequence in far less time than the usual 4-5 years. This enables them to go beyond calculus in their mathematics coursework prior to completing high school (Stanley, 1979, pp. 169-185).

In a typical secondary school mathematics course, mathematically gifted students are able to learn the material at a far faster rate than their classmates. Furthermore, the brilliant student often knows many concepts through previous learning prior to the introduction of the concepts in class. The relatively slow pace and frequency of repetition may cause the students to become bored and frustrated, which eventually may lead them to curtail their study of mathematics. There are numerous ways for schools to cope with this problem, such as special classes or supplemental instruction. The chief characteristics of appropriate opportunities are that the pace of instruction is quicker and the mathematics is more challenging than in the usual classroom. The challenge can be created by the pace of instruction itself, by including more difficult concepts and more depth in the course, or, as is usually the case, by a combination of added depth and an accelerated pace.

Regardless of which teaching strategy is used with the mathematically gifted, the problem of repetition of previously learned topics should be addressed. Whereas diagnostic testing is used in remedial instruction to pinpoint areas that need special attention, it is used with talented students to reveal areas of strength in a particular mathematics course. Already learned topics then form the basis for an instructional program that rapidly covers material not previously mastered. The model of diagnostic testing followed by prescriptive instruction has resulted in greater learning by mathematically talented students than usually occurs with traditional teaching techniques.

Implementation of the DT+ PI Model

The One-to-One Tutorial

Suppose that a teacher is to work individually with a certain mathematically talented youth who has never formally taken a course called Algebra I. There is a good chance that this bright student has learned a significant amount of first-year algebra, possibly through self-directed study. The teacher discovers what topics are known by administering a diagnostic test of

Algebra I. After the test, the teacher carefully examines the individual's answers to determine what concepts have and have not been mastered. If there are essentially no concepts that are found to have been learned, the whole of Algebra I should be covered in accordance with the usual sequencing of the material. If some concepts are found to be known, the teacher is able to accurately design a program of prescriptive instruction covering only those topics shown to be in deficit. It also may happen that the youth knows virtually all of Algebra I despite never having had a formal course in the subject. In that case, the teacher should quickly cover any remaining "holes" in the individual's knowledge of Algebra I, and then administer a diagnostic test of the next subject in the course sequence. The teacher follows the testing procedure until the course level at which the individual student needs to have more instruction than just filling in a few holes has been determined.

When the instructor has documented that Algebra I is completely mastered, the whole process starts again at the next level course. In subsequent courses, however, it may not be worthwhile to do any diagnostic testing. For example, an individual starting trigonometry would know very little about the subject through previous experiences unless there had actually been instruction in trigonometry as part of another course such as Algebra II. Diagnostic testing in this case would do little more than frustrate the student, as well as waste time in class. A few simple questions prior to testing can often indicate whether a diagnostic test would be appropriate.

A slight variation of this model is to reduce the amount of material covered by the diagnostic test to less than a whole course. For instance, the teacher may want to cover an Algebra I textbook one chapter at a time. Many books have chapter tests that are suitable for diagnostic purposes. Testing one chapter at a time has the benefit of thoroughly covering the course. There is, however, the danger of doing too much testing, thereby using up a great deal of valuable contact time with the student.

There are several principles in the instructional phase that are essential in making the DT $\rightarrow$ PI method as effective as possible for mathematically gifted youths. Of course, the foremost consideration is that the gifted student cover only those topics not already well understood. The pace at which instruction takes place is also crucial. The learning rate of talented youths is astounding when compared with that of the average person. Nevertheless, many teachers have a tendency to maintain preconceived notions of individuals' learning rates. An instructor may also try to do too much for a student, a habit understandably formed by teaching less able students. More than anything else, despite the obvious enjoyment of working with highly talented and motivated students, instructors of brilliant pupils need to discern when to step aside and allow their students to progress at a rate commensurate with the students' own ability levels. This does not imply that instruction is student-paced; on the contrary, it is up to the teacher to stimulate the student to move at a pace that is truly challenging.

A key element in effective instruction at any level is that there be continuous progress through the material to be learned. Homework and drill exercises should be assigned judiciously and not in such a quantity that they are overly repetitive and burdensome. Good teaching techniques in any situation involve an avoidance of "spoon feeding" students. This is especially important

with bright youths. Often they have been enrolled in classes taught below their capabilities, and consequently they have poorly developed study skills. Given the opportunity, such students will continue to rely on others for explanations, when confronted with difficult material, rather than on stretching their own minds through the use of their own faculties. The privilege of witnessing the enthusiasm and spark resulting from such independent and continuous progress, as well as the benefits to the individual of greater learning and self-reliance, more than compensate for any feelings on the instructor's part of not doing enough for a brilliant student.

A Class for Gifted Students

The DT $\rightarrow$ PI method of instruction is a valuable strategy for tutorial situations where a student meets separately with a teacher on a one-to-one basis. It has also been found to be an effective strategy in teaching classes composed entirely of mathematically gifted youths (Bartkovich and Mezynski, 1981.) In a classroom, initially the instructor goes through the diagnostic testing phase with each student. An individualized program of prescriptive instruction is then designed for each student. In practice, there are usually small groups of students in the class who receive a similar sequence of instruction. During class time, the instructor rarely lectures to the entire class. The pupils work through the prescribed instructional materials on their own or in small groups, requesting help from the teacher when necessary. The students may be at different points in their mathematics coursework and progress at different rates. The teacher uses class time to assist students individually or in small groups by answering questions, enriching various topics, and explaining difficult concepts. For example, the instructor may explain a concept to one student, assign a few problems to verify the individual's understanding, assist another student, and return later to check the first person's work.

It is important that the DT $\rightarrow$ PI model remain teacher-paced and not be self-paced by the class members. This is more difficult in a classroom, as compared to a tutorial situation. It requires excellent organizational skills on the instructor's part to monitor the progress of each individual and to be confident continuous progress is being accomplished. Rather than expecting students to progress as quickly as possible through the entire curriculum, the teacher must provide specific short-term objectives. Possible short-term goals include a deadline for completion of a certain section, the date by which a quiz must be taken, and a date to collect a homework assignment. The short-term objectives will vary from person to person and class to class. The task of the instructor is to discern the time required for an individual to complete a segment of a course. The teacher must continually encourage students to move at rates consistent with the students' individual knowledge and ability levels. It is extremely challenging for the teacher to monitor the progress of each individual as well as to firmly guide each student at the proper pace. Since the students are studying different topics, it is also no mean task for the teacher to capably answer, at any time, a question dealing with virtually any topic in a wide range of material.

Although the teacher assists and paces them, the students bear much of the responsibility for teaching themselves the course content. Instructional materials must be excellently and clearly written. The rationale of the DT $\rightarrow$ PI

model applied to the mathematically gifted is that these youths are able to learn from a high quality textbook without a teacher's explanation of each topic. By freeing them from the constraints of a lecture-style class, the more brilliant students are enabled to move ahead without waiting for the rest of the group. Likewise, within certain limitations, the individual who must study a specific topic further may do so without pressure to stay with the rest of the class.

The DT $\rightarrow$ PI Model at Duke

General Description

The Talent Identification Program (TIP) at Duke University utilizes the DT $\rightarrow$ PI model in the mathematics classes that are part of the Summer Residential Program for Verbally and Mathematically Precocious Youths. During the summer of 1981, brilliant youths from 30 states came to Duke for a three week program to study intensively mathematics or one subject from the humanities. There were 118 students in the six mathematics classes. The participants were mostly in the seventh grade or of seventh-grade age. Some had skipped one or more grades and were therefore of seventh-grade age but in a higher grade. Some were younger than the usual seventh graders.

The goal of the mathematics classes was for each of the 118 participants to learn as much precalculus mathematics as possible without sacrificing thoroughness or rigor. The courses included in precalculus were Algebra I, Algebra II, plane geometry, college algebra, trigonometry, and analytic geometry. The sequence in which students progressed from one course to the next depended on the sequence followed by their home school systems. This variability was made possible by the individualized nature of the instruction. The content of the courses compares favorably with the mathematics curriculum of any secondary school system. The textbooks used are widely known and are written by excellent, veteran authors. Evaluation was done with quizzes administered at specific points in each course.

On the first day of classes, all mathematics students took one or more diagnostic tests. They were given the test in the course at the next level in the precalculus mathematics sequence beyond what they had completed in school. Those who had not studied any precalculus courses took the Algebra I test. The results of the tests were used to design an individualized instructional program for each participant. The students were separated into somewhat homogeneous classes of about 20 each. Each class was taught by an instructor and one or two teaching assistants. The participants attended class for six hours Monday through Friday and three hours on Saturday for a total of 96 hours of contact time with the instructional staff. A voluntary study hall in the evenings was available for those who desired to do additional work, wanted to obtain special help, or who had not completed daily assignments.

The participants were always cognizant of what they needed to cover in the textbooks, as well as the sequence in which they were to proceed. They also were informed of where in a course they were to take quizzes. The number of problems to be solved was left flexible and largely up to a student's own discretion. After an experience with taking one quiz, the students quickly learned the level of mastery that was expected of them.

To enroll for the fast-paced mathematics classes, participants had to score at least 550 on the mathematics portion of the Scholastic Aptitude Test (SAT-M). An additional ten students were allowed to participate in the fast-paced mathematics classes based on scores of at least 500 on the SAT-M and 1000 total on the mathematics plus verbal portions of the SAT. For those enrolled in the mathematics classes, the average score on the SAT-M was 614, which is better than the scores obtained by 87% of the college-bound seniors who take the test. The group's mean score of 465 on SAT-V exceeds the scores of 54% of the college-bound seniors. An important point to emphasize is that these youths obtained their elevated scores at age 12 or 13, which is 4-5 years younger than most high school seniors.

Results

The achievements in terms of precalculus courses completed by the 118 participants are most impressive. On the average, 1.47 courses were completed by each student during the program. Contained in Table 1 is the number of courses completed by each participant. All but one student completed at least one course, and two exceptional individuals each finished four courses. At least 15 students were well prepared to enroll in rigorous calculus courses at their schools during the following year because of the quantity of mathematics they covered in the Duke program. Table 2 provides a breakdown of the number of students completing each of the precalculus courses. The number completing each course illustrates the diversity of the group in terms of precalculus knowledge. The numbers in Tables 1 and 2 do not include courses for which students were tested and placed out of without any instruction.

Table 1
Number of Courses Completed by
Students During 1981 Summer Program

Number of Courses	Number of students (N=118)
0	1
1	74
2	32
3	9
4	2

Table 2
Number of Students Completing Various
Courses During 1981 Summer Program

Course	Number of Students (N=118)
Algebra I	56
Algebra II	56
Plane Geometry	16
College Algebra	16
Trigonometry	20
Analytic Geometry	9

The accomplishments of the participants in these programs are further highlighted when compared to the typical student in secondary school. To complete a single precalculus course such as Algebra I, a student is required to attend approximately 180 classes, each 45-50 minutes in duration, over the course of a school year. Also consider that the same routine continues for 4-5 years to complete the entire precalculus sequence. Furthermore, in many school systems the precalculus sequence is not initiated until eighth or ninth grade, thus delaying the completion of all the precalculus courses until eleventh or, more likely, twelfth grade, even for the highly talented students. Contrast that with the tremendous leaps made by the summer program participants in the period of a few weeks. Clearly, the instructional staff did not actually teach the participants all of the mathematics for which the students received certification. By limiting teaching to only unlearned topics, the instructional staff made it possible for students to quickly complete subjects that would be drudgery if taken in school at a regular pace.

The Predictive Value of the SAT

An interesting outcome of the TIP 1981 summer mathematics program is the role of the SAT as a predictor of achievement. Bartkovich and Mezynski (1981) found that in individualized, fast-paced mathematics programs at The Johns Hopkins University, SAT-M had a statistically significant correlation with the number of courses completed, whereas SAT-V was not significantly correlated. The insignificant correlation of SAT-V and number of courses may have been due to the restricted range of SAT scores, since it was necessary for individuals to score high on both SAT-M and SAT-V to be eligible to enroll. Most of the students in the 1981 TIP classes were allowed to participate based solely on a high SAT-M score (at least 550), thus allowing SAT-V to vary freely.

Using only the data from the group of 108 that achieved an SAT-M score of at least 550 and not including the ten who enrolled through other criteria, correlation and linear regression analyses were performed. The correlation of the number of courses completed and SAT-M was 0.51 ($p < .001$); the correlation of courses learned and SAT-V was 0.17 (not significant). A linear regression analysis modeling the number of courses learned as the dependent variable and the SAT-M and SAT-V variables as predictors yielded regression coefficients of 0.062 and 0.005 for SAT-M and SAT-V, respectively. The F statistic had a value of 18.73 ($p < .001$). The multiple regression coefficient, R, had a value of 0.51, indicating that the regression equation accounted for 26% of the variance in the number of courses learned.

The implication of these statistical analyses is that SAT-M is an important predictor of achievement in fast-paced mathematics classes. The SAT-V is not found to relate significantly to achievement. When both SAT-M and SAT-V are used together to predict achievement, the SAT-V makes relatively little contribution. Variance in mathematical aptitude accounts for virtually all of the variance in courses learned; differences in verbal aptitude seem to have little direct effect on achievement in mathematics. This agrees with the previous findings by Bartkovich and Mezynski (1981). These conclusions must be interpreted in the context of the characteristics of the mathematics students. They all possessed high ability in mathematics; therefore, it is conceivable that when mathematical aptitude is above a certain level, verbal aptitude exerts negligible influence on the mathematics learning rate. Other factors

may become much more important, such as motivation, expectations, study habits, and discipline.

Implications

There are several important strengths inherent in the DT $\rightarrow$ PI model. It provides motivation, challenge, lack of repetition, and continuous advancement. The students develop greater self-reliance as they learn to teach themselves, a skill that has immeasurable, long-term benefits. The informal atmosphere and the principle of continuous progress make the learning enjoyable for most, especially compared with the relatively slow-paced mathematics courses to which they are accustomed.

A major weakness of the DT $\rightarrow$ PI strategy as applied to the mathematically gifted is the inadequacy of the standardized tests used for course certification. There is no recently created and developed series of standardized tests of the precalculus subjects. There are isolated tests that are of superior quality, but there is a need for an excellent series of current, reliable, and valid standardized tests to accompany the present precalculus curriculum.

In addition to its application with special classes for the mathematically precocious, the DT $\rightarrow$ PI model is useful in regular classes with the occasional brilliant student. In rural areas especially, the highly talented student is often the outsider in mathematics classes. The use of diagnostic testing provides a structure for such individuals to progress independently of the class, rather than moving along with less able classmates at a rate that is too slow to maintain interest and motivation. The teacher can guide a bright student by assigning cleverly designed problems and meeting periodically to discuss the subject material. The greatest need for independent work for the gifted students is in lower-level mathematics courses. In secondary school, as youths advance through the precalculus mathematics sequence, the less-able students tend to not elect the more advanced courses such as college algebra and trigonometry, which makes these courses more challenging for gifted students.

Conclusion

Mathematically gifted students are capable of benefitting from diagnostic and prescriptive programs of mathematics as much as any group in our schools. By virtue of their brilliance, these pupils often understand many topics before covering them in a formal classroom. These youths need special educational opportunities to stem the waste of talent brought about by boredom, frustration, and disillusionment with the standard curriculum, which is designed for average students. Diagnostic tools are an excellent aid in the design of an instructional program for highly talented students, including options such as special classes and independent study in regular classes. Summer programs such as those sponsored by Duke University's Talent Identification Program also have greatly benefited from the use of the DT $\rightarrow$ PI model in individualized classes with highly able mathematics students. The adaptability of the DT $\rightarrow$ PI model makes it feasible for schools to provide challenging instruction in situations where it is not viable to establish a special class, especially for schools that are small and have only a few mathematically gifted students. The DT $\rightarrow$ PI strategy is a way for teachers to provide for the needs of the occasional pupil

who can work at a level that is above the rest of the class.

The tremendous progress of the exceptionally able participants at Duke leads to the question, "What about students who are a bit less capable, but gifted nonetheless?" It seems that for many youths there is a middle ground between advancing to the same degree as the summer program participants and being in the regular classroom. If students in the upper one percent of their age group are able to complete precalculus mathematics in a fraction of the usual time, then surely individuals who are, for example, in the upper five percent could progress more rapidly than they do in typical secondary school mathematics classes. The DT $\rightarrow$ PI model provides an efficient, cost-effective way for schools to modify the mathematics curriculum for these students. Based on the alternatives presented here, curriculum design for the mathematically gifted need not be costly. Indeed, special opportunities must be made available if these individuals are to make maximum use of their potential.

References

- Bartkovich, K.G., and George, W.C. Teaching the gifted and talented in the mathematics classroom. Washington, D.C.: National Education Association, 1980.
- Kartkovich, K.G., and Mezynski, K. Fast-paced precalculus mathematics for talented junior high students: Two recent SMPY programs. Gifted Child Quarterly, 1981, 25, (2), 73-80.
- Conant, J.B. The American high school today. New York: McGraw-Hill, 1959.
- Stanley, J.C. Special fast-mathematics classes taught by college professors to fourth-through twelfth-graders. In D.P. Keating, ed. Intellectual talent: Research and development. Baltimore: The Johns Hopkins University Press, 1976, 132-159.
- Stanley, J.C. Rationale of the SMPY during its first five years of promoting education acceleration. In J.C. Stanley, W.C. George, and C.H. Solano, eds. The Gifted and the creative: A fifty-year perspective. Baltimore: The Johns Hopkins University Press, 1977, 75-112.
- Stanley, J.C. The study and facilitation of talent for mathematics. In A.H. Passow, ed. The gifted and talented: Their education and development. 78th Yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press, 1979, 169-185.

Linguistic Aspects of More/Less and Dimensional Adjectives Relevant to Diagnosis and Instruction

Carol Novillis Larson, University of Arizona

Some of the words investigated in the current psycholinguistics literature are more/less, big/small, long/short, tall/short, high/low, thick/thin, wide/narrow, and deep/shallow. These also happen to be terms that are very important in the mathematical lexicon since they are related to beginning number and measurement concepts. These terms are used when making many mathematical comparisons. Wales and Campbell (1970) observe that comparison is a very general feature of cognitive activity and is crucial to success on Piagetian tasks. Comparison has a linguistic aspect as well as a cognitive one. It is important for the researcher of children's cognitive structures not to assume that the child has the linguistic ability to comprehend and use terms such as more/less and the dimensional adjectives simple "because he uses these expressions in certain 'spontaneous' situations" (p. 373). This same point is relevant to the teacher of young children as well as to the researcher.

The results of this research illustrate that children do not have adult meanings for words; thus the usual assumption made when two adults are speaking cannot be made when one of the parties is a young child. This assumption is that both speaker and listener have approximately the same meanings for the words used. These studies show that it sometimes takes years before children's first meanings of words develop into complete adult meanings.

The psycholinguists are primarily interested in the linguistic characteristics of the terms investigated and in the development of a theory of language acquisition. But their analysis of the meanings of the various terms and their discovery of pertinent features of research tasks that produce different rates of success in the children tested can also provide valuable information to both teachers and diagnosticians working with young children. Task analysis is used in many diagnostic models (Underhill, Uprichard, and Heddens, 1980). Taking a task and analyzing it in terms of prerequisite subordinate tasks is usually necessary before writing a diagnostic test or planning an instructional program. This new psycholinguistic information can be of help in the analysis of tasks for the specific concepts related to the terms discussed in this paper.

The purpose of this paper is to review some of the psycholinguistic literature on young children's meanings of more/less and the following dimensional adjectives: big/small, long/short, tall/short, high/low, thick/thin, wide/narrow, and deep/shallow. The review will be organized by the terms being investigated: more/less and the dimensional adjectives. Each section will have three subdivisions: (1) a discussion of the meanings of the terms and other semantic characteristics that have been explored as being relevant to children's acquisition of these terms; (2) the results of some of the research on children's comprehension of these terms; and (3) implications for diagnosing and teaching.

More/Less

Semantic Characteristics of More/Less

Researchers who have investigated children's meaning of more and less have

attempted to describe the adult meanings of these terms. They have also identified what appear to be some semantic characteristics of these terms that might be related to the words' order of acquisition.

More and less are the comparative forms of the adjectives much/many and little; they can be used in comparisons of both discrete and continuous quantities. More also has an "additive" sense as well as a "comparative" sense; less has only a "comparative" sense. One can say 'give me some more' but not 'Give me some less'.

Oppositeness of meaning is an important semantic relation between terms. More and less are antonyms according to Lyons' (1968) description of three types of oppositeness: complementarity, antonymy, and converseness. In pairs of antonyms, one term is usually considered to be semantically "marked" with respect to the other term in the pair. The "unmarked" term is the one that can be used in a neutral as well as a contrastive sense. For the pair much/little, much is unmarked and little is marked. In the question "How much money do you have?" much has a neutral meaning referring to amount in general. Whereas the question "How little money do you have?" is contrastive: the implication is that you have little money (H. Clark, 1970).

Markedness is often related to another linguistic characteristic, polarity. Lyons (1968) discusses this relationship for certain antonyms:

The fact that the distinction between antonyms is neutralized in certain syntactic positions contributes, no doubt, to our feeling that one antonym has a 'positive', and the other a 'negative' polarity. We tend to say that small things 'lack size' rather than large things 'lack smallness'. And, in general, the 'unmarked' antonym is used for what is felt as 'more than', rather than 'less than', the norm (p. 476).

Markedness and polarity play a very important role in the interpretation of the research to be presented. E. V. Clark (1973a, 1973b, 1975, 1977) has predicted that children will first acquire the meaning of the positive, unmarked term in a pair of antonyms. She further predicts that when children first begin to use the marked, negative term they treat it as a synonym of the unmarked term.

One theory of linguistic meaning is based on the notion that each word is a collection of basic elements of meaning called "semantic markers" or "semantic features" (See Katz & Fodor, 1964). In this theory antonyms share all semantic features except for the contrastive value of polarity. In order to differentiate between antonyms the child needs to attend to this one difference.

The adjectives more/less are used in many diverse situations. Identifying these situations has been important in order to design research tasks that will help to uncover children's meanings of these terms. Donaldson and Balfour (1968) offered the following analysis:

1. More/less may apply to discrete units or to continuous quantities.
2. More/less may be used to compare one entity across time, two entities at a given point in time, or two entities across time.

3. When comparing two entities across time, various types of changes may or may not take place. When change occurs, either one or both of the entities may change. If both change, both may increase or decrease, or one may increase and the other may decrease. In the latter case the changes may be related or independent. They are related if a transfer is made from one entity to the other. The results of a change may or may not require a change in the initial judgment. The directions of the change may or may not be in accord with the final state of the two entities relative to one another.
4. When change occurs the person making the comparative judgment may or may not make the change himself/herself.

Other possibilities for change involve changes in irrelevant variables coupled with and without changes in quantity.

This analysis can be applied to designing diverse teaching activities that could help children extend their concepts of more and less in mathematical situations, as well as to learn the meanings of the terms.

Research on Children's Comprehension of More/Less

One of the most consistent results in the research of children's comprehension of the terms more/less is that young children learn the comparative meaning of more prior to that of less (Donaldson & Baifour, 1968; Palermo, 1973, 1974; Weiner, 1974; Wannemacher & Ryan, 1978). By five years old, most children are responding correctly to tasks that require the comprehension of more. Palermo's (1973) results with children in kindergarten through second grade illustrate the difference in acquisition rate of the term less. All children in this study responded correctly to the more questions; the frequency of errors for each grade level on the less questions is presented in Table 1.

Table 1
Frequency of Errors on Less Tasks

Number of Errors on 19 Tasks	Number of Subjects at Each Grade Level		
	K	1	2
15-19	25	9	6
3-14	4	3	1
1-2	3	20	25
Total Subjects	32	32	32

In some of the cited studies children responded to less as if it meant more and the researchers therefore claimed that the children's meaning for less was that it was a synonym for more (Donaldson & Balfour, 1968; Donaldson & Wales, 1970; Palermo, 1973, 1974). The children in these studies were told to choose between only two situations; so, a child either made the correct choice for less or was forced to choose the more response. The research task does not permit any other pattern of incorrect responses. Other data conflict with the theory that children's first meaning for less is as a synonym for more (Weiner, 1974; Wannemacher & Ryan, 1978).

Wannemacher & Ryan (1978) report that less was easier for children in their study than in many of the other research studies. This difference in results may be related to the fact that this study differed from previous studies in that (1) children were given an explicit standard, and (2) more and less were always tested on the same trial. Donaldson and Balfour (1968) tested more and less on separate days; Palermo (1973) tested more and less in separate trial blocks; and Palermo (1974) tested only less. Other studies that report young children having a good understanding of less also tested more and less on the same trial (Beilin, 1965; Griffiths, Shantz & Sigel, 1967).

Holland and Palermo (1975) lend support to the fact that using more and less together seems to facilitate children's comprehension of less. Sixty-two 5-year-olds who did not know the meaning of less were assigned to a control or to an experimental group. In the training session the children in the experimental group were repeatedly asked to make more and less distinctions for the same pair of stimuli. All but three of these children succeeded in learning the meaning of less, whereas only two children in the control group improved their performance on the less post-test.

The surprising outcome of the training session was the speed and ease with which children learned the appropriate reference for "less": most were responding correctly and unhesitatingly to "less" questions by the end of the first part of training and were then able to answer questions and solve problems of more difficult tasks without further training (p.441).

The responses of children in contexts when possible answers are not given (unrestricted contexts) are different from those in context when the response alternatives are provided (restricted contexts). Wannemacher and Ryan (1978) report that in restricted contexts all 3- and 4-year-olds selected a response with apparent conviction, but in the unrestricted contexts they either responded correctly, answered "I don't know" or made no response. If only restricted tasks had been used, the researchers would have concluded that 4-year-olds comprehended the meaning of less as well as did the 5-year-olds. But the responses of 4-year-olds on the unrestricted tasks made it clear that they were at some intermediate level of competence between knowing the meaning of less in all contexts or in none. Wannemacher and Ryan stated the ultimate conclusion of their study as follows:

The way young children interpret a word in a given situation reflects not only their linguistic knowledge, but also the operation of response biases, or nonlinguistic strategies elicited by particular task contexts and experimental procedures (p. 667).

99

This conclusion seems to apply to all of the studies discussed in this section and is extremely important in applying the research results and conclusions to new and novel situations.

Further evidence of this was reported by Donaldson and McGarrigle (1974). In investigating children's comprehension of more and all, they point out that children shifted the basis for their judgments when confronted with different tasks. When 4-year-olds were shown two rows of cars, 34 out of 40 children chose the larger set when asked which row had more. When two garage structures, one with space for six cars and one with space for four cars, were positioned over the rows of five and four cars respectively, 13 children changed their response to more by choosing the set of four because the garage was full. An example of the children's reasoning was expressed by one child who said, "There's more on that shelf (pointing to the smaller subset) because there's enough to go in there" (p. 189).

When 11 of the 86 children in Weiner's (1974) study responded as if less meant more, she attributed this to a nonlinguistic preference for the longer row. A follow-up study by Weiner (1974) testing children's ability to make quantitative comparisons when the terms more/less were not used lends support to this observation.

E. V. Clark's (1973a) results in a series of studies of children's meanings of in, on and under support the hypothesis that children first rely on non-linguistic strategies in their acquisition of these terms. In applying this observation to the research on more/less, E. V. Clark (1973a) interprets the results to mean that at first children have an incomplete meaning for both more and less. They probably know that both words refer to amount; but they also have a nonlinguistic strategy of usually choosing the greater of two amounts. This would account for the children appearing to know the meaning of more and in seeming to treat less as a synonym for more. Some of Donaldson and Wale's (1970) reports of individual children's responses to some tasks lend support to the hypothesis that some children are using more and less to mean simply amount. Some children's replies to the question "Does one tree have more apples on it than the other?" were "Both of them, that one does and that one, both the trees, they two ones, and each tree" (p. 248).

Implications for Teaching and Diagnosing

Research on children's acquisition of more/less can be applied to teaching at two levels: the general and the specific. The general level will be discussed in the last section of this paper. In this section some specific strategies for teaching the meaning of the terms more and less will be briefly outlined.

When teaching the meaning of the terms more/less, the teacher needs to recognize that less is more difficult for young children than more: i.e., its meaning is learned later in the natural order of language acquisition. So the teacher must not expect children to know the meaning of less just because they know the meaning of more. More is used more frequently by adults than is less. In a comparative situation adults will usually express the quantitative relationship by using the term more unless there is a special reason for expressing the quantitative relationship in the opposite direction. Teachers will also unconsciously do this unless they have a special reason not to, but they do have a reason--their special reason is to help children learn the meaning for less. This requires conscious awareness of the aforementioned facts and a definite, well-thought-out sequence of instruction.

The research discussed in the previous section indicates what some facets of a sequence of instruction might be:

1. Teach more and less together.
2. At first, use a standard and arrange objects in rows that facilitate one-to-one correspondence.
3. Use a large variety of activities and materials.
4. Some of the activities can be restricted (forced-choice), but many of the activities should be unrestricted (open-ended).
5. Activities should not involve just static comparisons; some of the activities should involve comparing entities when changes are made similar to those listed previously.

Similar considerations should also be taken into account when designing tasks for testing children's comprehension of more and less. For, as the research shows, many of children's meanings for words are context specific and an overall general meaning seems to take time to develop. Therefore, prior to deciding that a child has a full adult meaning for more or less, the teacher needs to test the child on diverse tasks that vary in context.

Dimensional Adjectives

Semantic Characteristics of the Dimensional Adjectives

The dimensional adjectives, also called spatial adjectives, to be discussed in this paper are big/little (or small), long/short, tall/short, high/low, wide/narrow, deep/shallow, and thick/thin. The dimensional adjectives comprise a semantic field. Terms that are members of the same semantic field share some of the same semantic features, i.e., terms in a semantic field are related to each other in meaning. The semantic feature (Spatial Extent) is shared by all terms in this semantic field. Other semantic features shared by subsets of this field are (+Polar). The dimensional adjectives come in antonym pairs that differ in polarity. The first term listed in each pair has the semantic feature (+ Polar) and the second term listed in each pair has the semantic feature (- Polar). All of these adjectives have a comparative form and a superlative form marked by the morphemes -er and -est respectively.

Each pair of dimensional terms is associated with one of the following dimensional scales: size, length, tallness, height, width, depth, and thickness. Except for the pair big/little, the first member of each pair is used as the basis for each scale name. Each scale is measured from some primary reference or zero point. The first element in each pair labels the positive direction away from the primary reference point and hence can be used in a neutral sense in discussing the dimension. For example, in "How long is that board?", long is used in a neutral sense indicating the scale of length. When used in a contrastive sense, the positive term indicates that the object being discussed is greater in extent than a standard for that particular class of objects. This standard, that is particular for each class, determines a secondary point of reference on each scale. In the sentence "That is a long board.", long indicates that the length of the board is greater in extent than some standard board, i.e., the board extends beyond the secondary standard point. The negative term is

only used to measure from the standard point back to the zero point; hence the negative term cannot neutralize to extend to the whole scale. In "That is a short board.", short indicates that the length of the board is less than the standard board. The use of the positive term in measure phrases further demonstrates the neutral use of the positive term; "two hours long" can be used, but not "two hours short". Because the positive term can at times neutralize to denote the whole scale, it is considered to be "unmarked" and the negative term is considered to be "marked" (cf. Greenberg, 1966; Bierswisch, 1967; H. Clark, 1970, 1973).

The terms big and little (small) differ from the other antonym pairs in that they are more general in meaning, i.e., fewer restrictions are placed on their use; they can be used to refer to one-, two-, and three-dimensional objects. Hence they can be used in many circumstances that overlap the usage of the other adjectives in the semantic field. In some sense big/little are superordinate to the other adjectives in this field. "The dimension of comparison underlying big and little is spatial extent alone; all other spatial relatives require further features for their specification, such as (vertical)" (Carey, 1978, p. 277). Therefore, all of the other terms are more specialized in meaning: they have more semantic features and restrictions for their use, and as a result are considered semantically more complex. Bierswisch (1967) has identified other semantic features that further differentiate between terms in the field of dimensional adjectives. Even though all dimensional adjectives share the semantic feature (Spatial Extent), big/little apply to (n-Space) whereas the other adjectives in this field refer to only one particular dimension (l-space). This one dimension can be further specified in terms of (+ Vertical). Also, when describing two- and three-dimensional objects, some dimensions are not used to name the most extended dimension of the object; these have the added semantic feature of (+ Secondary). In terms of this analysis, the following are some examples: long/short have the semantic features (l-Space), (- Vertical); and wide/narrow have the semantic features (l-Space, (- Vertical), and (+ Secondary) (E. Clark, 1972).

Research on Children's Meanings of Dimensional Adjectives

Many psychologists have chosen to study children's acquisition of dimensional adjectives because these words form a semantic field--"in the linguistic system of adults these adjectives form a closely-related and highly structured set" (Wales & Campbell, 1970, p. 374). Psycholinguists are interested in answering questions such as "Which terms are acquired first?" and "What are the linguistic characteristics of the terms, e.g., semantic complexity and polarity, that cause some terms to be acquired before others?"

Children appear to know that the dimensional adjectives are related. When they become aware of a new pair of dimensional adjectives, they seem to associate the positive term with big and the negative term with little or small (E.V. Clark, 1972). In accord with this, Wales and Campbell (1970) report that big and wee (wee is used in Scotland instead of little) were the adjectives most frequently used in a production task. The production task consisted of asking the children questions about a set of pictures that were placed in the context of a story.

The questions and pictures were designed to suggest specific dimensional adjectives, e.g., long/short. This result is also supported by Brewer and Stone's (1975) conclusion that polarity is acquired before dimensions are sorted. Brewer & Stone tested children's comprehension of the 12 specific dimensional adjectives by using a task containing objects representing two independent dimensions. This design was chosen to make possible a discrimination between children's knowledge of the dimensions being tested and their knowledge of polarity. The most common error made was to choose the exemplar for "the different dimension, same pole", rather than for "the same dimension, opposite pole". This difference was significant. In studies where all choices vary only on the dimension being tested this result would not be possible.

In all of the studies examined, children were consistent in the order in which they acquired the seven antonym pairs: 1) big/little; 2) tall/short, long/short; 3) high/low; 4) thick/thin; and 5) wide/narrow, deep/shallow (E.V. Clark, 1972; Eilers, Oller & Ellington, 1974; Brewer & Stone, 1975; Bartlett, 1976). This ordering corresponds directly to the relative semantic complexity of the terms as described in the previous section.

The positive adjective in each antonym pair is usually better known than the negative term and is learned first (Donaldson & Wales, 1970; Brewer & Stone, 1975). Some exceptions to this were reported by Bartlett (1976), and Eilers, Oller, and Ellington (1974). Bartlett (1976) reports no significant difference in children's comprehension of the positive and negative terms of four pairs of dimensional adjectives when single-dimension stimuli were used; but positive terms were significantly easier when double-dimension stimuli were used. Eilers, Oller and Ellington (1974) report that 2- and 3-year-olds made significantly more errors on positive terms (big, long, and wide), than on negative terms (little, short, and narrow). However, when size preference was controlled for, there was no significant difference in children's comprehension of positive and negative terms in each antonym pair.

In order to investigate why positive adjectives in antonym pairs were learned first by children, Klatzky, Clark and Macken (1973) designed a training study using three-letter nonsense words for each of the following eight terms: big/small, high/low, long/short, and thick/thin. The major result was that following training children were more successful in responding correctly to the positive nonsense words than to the negative ones. Since the children were given the same opportunity to learn both terms of a pair, the conclusion was that the pattern of acquisition was due to a conceptual basis and not to the fact that the positive terms are more frequently used by adults.

Implications for Teaching and Diagnosing

The terms big/little are understood by children at a very young age. When teaching the other dimensional adjectives the research indicates that using objects that vary on more than one dimension is more difficult for children--but that it is also necessary if the reference of the term being used is to be clear. When children discuss objects varying on only one dimension they usually use the more general term--big or little, rather than a specific term such as long or short--for in this case there really is no reason to be more specific. A good example of this is the case of using Cuisenaire rods to teach the terms long/short. Many teachers and students using these rods often refer to their overall size rather

than to their length. They often say "Which is bigger or larger?" rather than "Which is longer?"

If children are to sort out the various dimensions, then they need to be given materials that differ on two or three dimensions, e.g., a set of boxes of various sizes. Children can be asked to compare the length, width, and height of pairs of boxes and to seriate them by length, width, or height. These kinds of activities would help children differentiate the three dimensions and learn the terms associated with each one.

In doing activities like those just described, the teacher needs to consciously use the negative terms (short, narrow, and low) as well as the positive adjectives in each antonym pair. Fluent adult speakers of the language use the positive adjectives much more often than the negative ones. If children are to learn both terms at the same time then they must be given extra exposure to the use of the negative adjectives.

Discussion

Only one small aspect of children's language development has been examined in this paper--children's meanings for a few but important terms. Admittedly, there are many limitations to the research reported. This is due to the difficulty of trying to infer word meaning from children's responses to directions and questions, as well as to the problem of trying to get children to produce specific terms in contrived situations. Yet from this research several general trends are worth noting.

1. Children's first meanings for the terms reviewed were shown to be incomplete. Because these terms apply to many different contexts from simple to complex, there are times when children who have an incomplete meaning for a term appear to comprehend the term and use it correctly. So if knowledge of these terms is important in understanding the questions being asked in Piagetian tasks or in investigations of mathematical concept development, then it is not enough to test children's knowledge of these terms on just any comprehension task. One must ascertain children's meanings of the terms in situations requiring the same linguistic complexity as the cognitive task. This also applies to the diagnosis of children's mathematical knowledge.

2. A lack of full meaning for words exists for many other terms besides those discussed in this paper. As this research shows, with age and experience children do learn the full meaning of words for which they previously had partial meaning. An important point is that children are constantly being exposed to new words. So at any given time there exist words in their lexicons for which they have only partial meanings. This is true for adults as well as children.

3. Researchers, diagnosticians, and teachers need to become more sensitive to the phenomena of developing and expanding word meaning. We can infer from this research that mathematics instruction needs to be designed to encourage student discussion and exploration. The terms discussed in this paper are part of the vocabulary of the general populace and are not technical terms. With or without instruction, most children will eventually develop meaning for these terms.

However, other mathematical terms that children need to learn are not part of this general vocabulary. Hence, providing the opportunity for students to use technical mathematical terms in many diverse settings is important, if developing a generalizable meaning of these terms is a goal of school mathematics.

There is a growing body of research in the area of child language acquisition that needs to be considered when designing instruction, diagnostic tests, and interviews--especially when interpreting their results. This review of children's meanings of some terms that are part of the "language of mathematics" will hopefully help mathematics educators become more aware of one aspect of children's linguistic development. During this century mathematics educators have become aware of the importance of children's cognitive structures in their learning of mathematics. Mathematics curricula and diagnostic and remediation approaches have been designed based on learning theory and child development.

References

- Bartlett, E.J. Sizing things up: The acquisition of the meaning of dimensional adjectives. Journal of Child Language, 1976, 3, 205-219.
- Beilin, H. Learning and operational convergence in logical thought development. Journal of Experimental Child Psychology, 1965, 2, 317-339.
- Bierswisch, M. Some semantic universals of German adjectivals. Foundations of Language, 1967, 3, 1-36.
- Brewer, W.A. & Stone, J.B. Acquisition of spatial pairs. Journal of Experimental Child Psychology, 1975, 19, 299-307.
- Carey, S. The child as word learner. In M. Halle, J. Bresnan, & G.A. Miller (Eds.), Linguistic theory and psychological reality. Cambridge, Mass.: MIT Press, 1978.
- Clark, E.V. On the child's acquisition of antonyms in two semantic fields. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 750-758.
- Clark, E.V. Non-linguistic strategies and the acquisition of word meaning. Cognition, 1973, 2, 161-182. (a)
- Clark, E.V. What's in a word? On the child's acquisition of semantics in his first language. In T.E. Moore (Ed.), Cognitive development and the acquisition of language. New York: Academic Press, 1973. (b)
- Clark, E.V. Knowledge, context, and strategy in the acquisition of meaning. In D.P. Dato (ed.), Developmental psycholinguistics: theory and application. Washington, D.C.: Georgetown University Press, 1975.
- Clark, E.V. Strategies and the mapping problem in first language acquisition. In J. Macnamara (Ed.), Language learning and thought. New York: Academic Press, 1977.
- Clark, H.H. The primitive nature of children's relational concepts. In J.R. Hayes (Ed.), Cognition and the development of language. New York: John Wiley & Sons, Inc., 1970.
- Clark, H.H. Space, time, semantics, and the child. In T.E. Moore (Ed.), Cognitive development and the acquisition of language. New York: Academic Press, 1973.
- Donaldson, M. & Balfour, G. Less is more: a study of language comprehension in children. British Journal of Psychology, 1968, 59, 461-471.
- Donaldson, M. & McGarrigle, J. Some clues to the nature of semantic development. Journal of Child Language, 1974, 1, 185-194.

- Donaldson, M. & Wales, R. On the acquisition of some relational terms. In J.R. Hayes (Ed.), Cognition and the development of language. New York: John Wiley & Sons, Inc., 1970.
- Eilers, R.E., Oller, D.K., & Ellington, J. The acquisition of word meaning for dimensional adjectives: the long and short of it. Journal of Child Language, 1974, 1, 195-204.
- Greenberg, J.H. Language universals. Paris: Mouton & Co., 1966.
- Griffiths, J.A., Shantz, C.A., & Sigel, T.E. A methodological problem in conservation studies: the use of relational terms. Child Development, 1967, 38, 841-848.
- Holland, V.M. & Palermo, D.S. On learning "less": Language and cognitive development. Child Development, 1975, 46, 437-443.
- Katz, J. & Fodor, J. The structure of a semantic theory. In J. Fodor & J. Katz (Eds.), The structure of language. Englewood Cliffs, N.J.: Prentice-Hall, 1964.
- Klatzky, R.L., Clark, E.V., & Macken, M. Asymmetries in the acquisition of polar adjectives: linguistic or conceptual? Journal of Experimental Child Psychology, 1973, 16, 32-46.
- Lyons, J. Introduction to theoretical linguistics. Cambridge: University Press, 1968.
- Palermo, D.S. More about less: a study of language comprehension. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 211-221.
- Underhill, B., Uprichard, E., & Heddens, J. Diagnosing mathematical difficulties. Columbus, OH: Charles E. Merrill Publishing Co., 1980.
- Wales, R. & Campbell, R. On the development of comparison and the comparison of development. In G.B. Flores d'Arcais & W.J.M. Levelt (Eds.), Advances in psycholinguistics. New York: American Elsevier Publishing Co., Inc., 1970.
- Wannemacher, J.T. & Ryan, M.L. "Less" is not "more": a study of children's comprehension of "less" in various task contexts. Child Development, 1978, 49, 660-668.
- Weiner, S.L. On the development of more and less. Journal of Experimental Psychology, 1974, 17, 271-287.

Improving Computation--A Look at Estimation and Approximation As Significant Elements in Achieving Understanding and Correctness

Douglas H. Crawford, Queen's University, Kingston, Ontario

The area of estimation and approximation is of considerable significance, especially in business and industry, but also for the ordinary citizen. The importance of the area for the mathematics curriculum has been argued by a number of writers in recent years. As Payne and Seber (1959) stated:

Often in arithmetic and in more advanced mathematics we are concerned with exact computation but neglect an equally important aspect, estimation. We may develop the algorithms for multiplication with understanding and yet fail to challenge pupils to think about the other relationships between numbers. We teach that 5×7 is 35, but leave the child to his own devices in realising that 5 times a number more than 7 is a product more than 35. Though we teach him how to find the area of a rectangle and introduce the formula $A = lw$, we often overlook the highly useful aspect of having pupils estimate the length and width of a room to get a sensible estimate of the area of the rectangular floor. Without a design for learning which includes estimating and making sensible approximations, it is highly unlikely that students will achieve this by themselves. Certainly, estimating results of computations should become an integral part of the instructional program in arithmetic. (p. 183)

In reporting on estimation as an aspect of the National Assessment of Educational Progress, Carpenter (1976) noted that recent recommendations identifying basic computations essential for enlightened citizenship have all included estimation as a fundamental skill that every student needs to learn. He continued,

In fact, estimation is more important than precise calculation for many common uses of mathematics. (p. 296)

These pleas and recommendations have been heeded by the National Council of Teachers of Mathematics, which has included the skills of estimation and approximation as one of the ten basic skills in their manifesto An Agenda for Action (NCTM, 1980).

In this article, attention will be focused on an analysis of the significance of estimation and approximation with a sketch of a possible approach to their teaching.

The Concepts of Estimation and Approximation

What do we mean by these concepts? It is best to attempt an answer by using one or two illustrations.

I place a pile of saucers on the table and ask, "How many saucers are there?" Your reaction is probably to count them and answer, "Nine". Suppose I now show you a glass jar with toffee candies in it, and ask, "How many candies are there?",

but add "Guess, don't count"; then your answer will be an estimate. Your estimate may be, in this case, exactly right, but the chances are that it won't. How nearly correct it is now brings in the concept of approximation. A "good" estimate, for example 46 when the actual answer was 48, will mean a close approximation. The estimate is the actual number you gave as the answer. The approximation requires you to consider the relation between your estimate and the correct or exact number.

Estimation probably most often occurs in connection with measuring things. For example, when I ask you, "How far is it to Hershey from New York City?", you might reply, "About 100 miles." In this case we are making an estimate which involves not just a number but a unit. In this case, as Bright (1976) says, "estimating is the process of arriving at a measurement or measure without the aid of measuring tools." Bright goes on to describe eight basic kinds of "measurement estimation" and to discuss the preparation for and the teaching of these estimation skills. He concludes that this type of estimation--

- 1) can help solidify the concepts of number and counting as well as the concepts of more/less, larger/smaller, heavier/lighter and so on;
- 2) can help describe the world;
- 3) enables practice in computation. (p. 103)

In a scholarly presentation to the Third International Conference on Mathematics Education, Guilbaud (1976) discusses the relation between the exact and the approximate, and the need to see them in proper mathematical perspective. We should realise that "to be obliged to approximate and to know how to do it properly is not a weak point of mathematics but its strength and its glory," and think of an exact calculation as one for which no approximation is needed.

Guilbaud appeals to everyday life to show how essential an ingredient approximation is:

Just try yourself; take a newspaper. They are full of figures today: most of the figures are given implicitly or explicitly, with a qualification as to the degree of certainty. The population of a town, an industrial output, a distance, a temperature, a duration, a speed, a percentage, etc. How are we to find our way about? An approximate value, as common sense says, is that which is not exact. Is it a lie then? I would not deny that newspapers sometimes contain lies. But there is not always so much malice. No! Talking and thinking by means of "about", "nearly" is a necessity. The enquiry can be made more systematic by choosing special areas of human activity. What is the accuracy used in sports timing? How are unavoidable uncertainties expressed in geodesy? What is the accuracy of financial calculations? And in demography? In economics? With a little imagination, one can then prepare numerous study trips for mathematics education, because it is essential to know what the different trades do with the numerical categories they have instituted. (p. 126)

He goes on to identify four levels of increasing sophistication of "approximate" thinking:

- 1) an approximate value.
- 2) an interval.
- 3) an indefinite sequence of intervals.
- 4) a tolerance or confidence interval with an associated probability.

Flow Chart for Developing the Concepts of Estimation and Approximation
in the Early Grades

It is clear that an understanding of number itself is essential before estimating can be a meaningful process. Since an estimation may be also of a measurement, basic ideas about measurement, especially the understanding of applying a known unit (arbitrary or standard), and the fact that measurement is itself an approximate process will need to be integrated to achieve adequate understanding. The other most related concepts are: "less than" and "more than", "closeness", "betweenness", and "rounding". A flowchart showing one possible network relating the development of these concepts is shown in Figure 1.

Approaches to Teaching Skills in Estimation and Approximation

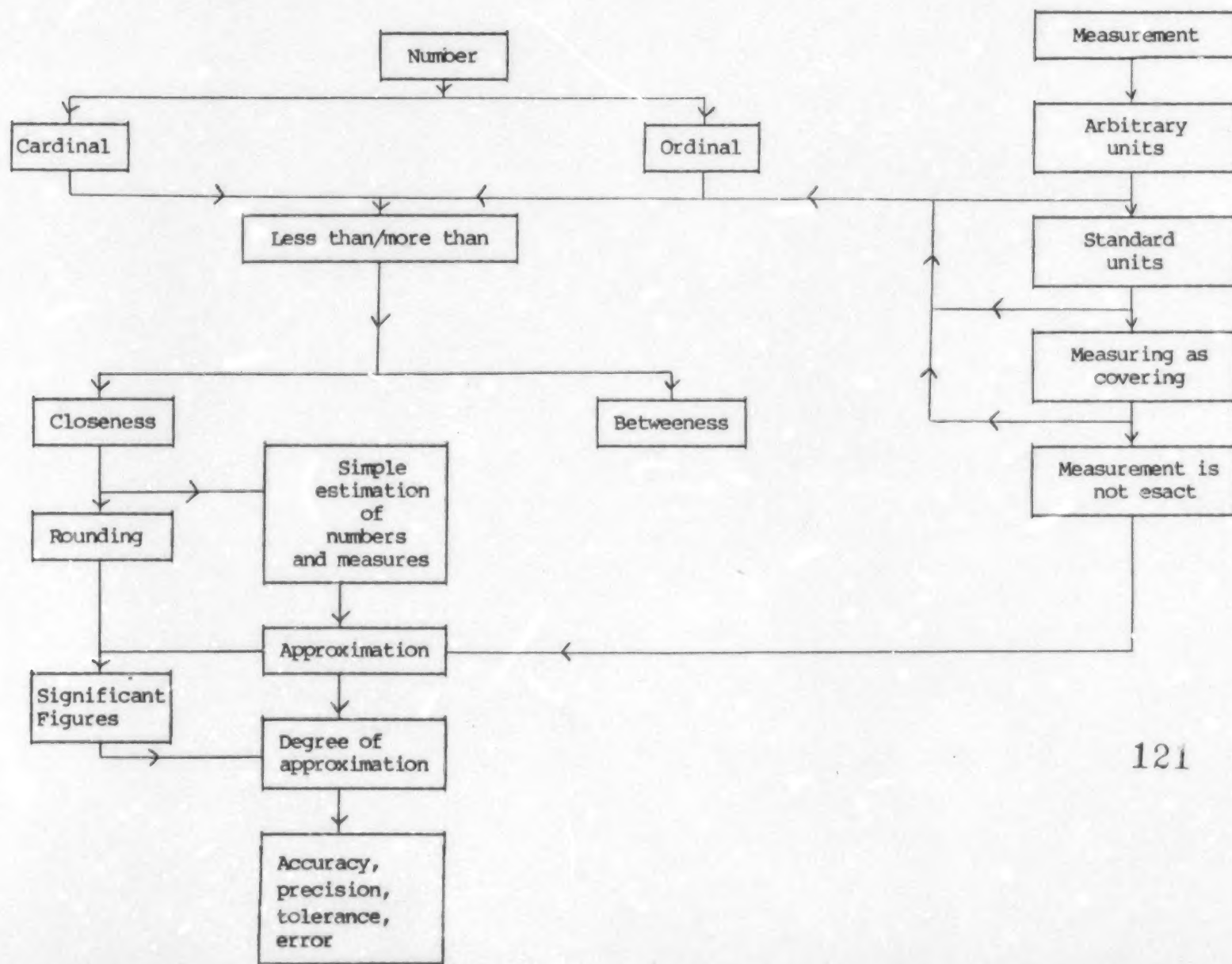
According to Trafton (1978), estimation and mental arithmetic skills should form part of the minimal competence in mathematics expected of all citizens. He cites a number of reasons, including:

- 1) Their value in many practical situations, e.g., in buying and traveling
- 2) Their bringing a new dimension and vitality to the study of computation.
- 3) The increase in reasoning, insight and command over numbers which they can bring.

Trafton goes on to outline four useful approaches to developing estimation skills. Two of these are somewhat similar--rounding numbers and then computing, and a variation of this when the "front-end" or largest part of the numbers is used. A third makes use of a reference point with which you compare some measure. For example, two articles may be priced at \$3.49 and \$2.75 and you have \$5.00 in your pocket. Can you buy both? Here the simple need is to determine if the sum of the two prices is more or less than \$5.00. This type of situation is very common. Finally, Trafton outlines a fourth estimation situation where the goal is to determine whether the answer is reasonable or not. In this case, what is needed is a range, or interval, or order of magnitude within which the answer must be, to check against the actual answer or simply to yield an acceptable estimate for some desired purpose. This type of estimation is particularly necessary with the widespread use of the hand-held calculator--as a check against mechanical or conceptual errors in using it.

Trafton also gives some useful suggestions for carrying out estimations mentally which use the "front-end" computing approach. For example, in computing 6×72 , proceed mentally by obtaining 6×70 and then adding 6×2 . He concludes by arguing for meaningful, systematic instruction in estimation and mental arithmetic skills.

Figure 1. A Flow Chart for Developing the Concepts of Estimation and Approximation



At a more practical and basic level, Carlow (c. 1969) has outlined a sequence of estimation exercises appropriate for the K-6 grade levels. They begin with estimating which is more or less, and conclude with estimating various ranges of lengths and distances. Carlow is concerned that a variety of estimates revealing approximate relationships should be made by the student prior to the telling of exact relationships. For example, a first approximation to the area of a circle is:

$$\underline{A} = 3\underline{r}^2 \text{ and a second } \underline{A} = \frac{22}{7}\underline{r}^2.$$

With a cleverly sequenced and imaginative program of estimation,... the student cannot help but become aware of number properties and relationships in the world around him. (p. 4)

Bitter and Mikesell (1980) state that in order to gain facility in estimating, understanding and skill must be developed in--

- a) whole-number and decimal place-value concepts,
- b) rounding numbers,
- c) working rapidly and accurately with multiples and powers of ten.

Learning to Estimate and Approximate

As already indicated, it is assumed here that the skills of estimation and approximation should be integrated into the elementary curriculum.

Thus, when one-digit numbers are being learned, their order relations should be learned too--initially at the concrete level and only then internalized to the abstract and symbolic level. For example, Bill can be given 6 beads and Jean 4. They are asked, "Who has more?" "How do you know?" They can match them and then answer the question how many more? Many simple exercises like this, including simple games with structured materials will fix the ideas and meaning of more, less, fewer--first at the concrete level, then at the symbolic levels. Eventually, students should be able to make the transition from the concrete situation to the symbolic one of say $3 < 4$ (see, for example, Fuys and Tischler, Chapter 2).

Children should be given opportunities to estimate numbers, using handfuls of marbles, chips, or buttons. When we try to guess "how many" without counting or measuring, we are giving an estimate. If the estimate is close to the actual number it is a good estimate. We then call it an approximation.

The next stage is to develop the related ideas of betweenness and closeness. An abbreviated version of one way to approach this is shown in Figure 2. We can then develop the idea and rules for rounding to the nearest ten, and extend them to three- and four- digit numbers, and to other levels of rounding. The student is then ready to apply these skills in computation situations.

Summary

This article has focused on the continuing importance of the concepts of estimation and approximation. With the advent of the calculator and the personal computer, meaningful computation will be even more dependent on these two concepts. By giving continuing serious attention to them in teacher education and in the classroom, the learning of basic arithmetic and its real-life applications can be significantly improved.

Figure 2. An Abbreviated Sequence for Learning to Estimate and Approximate a Two-Digit Number

- A.
 - 1. Make the number 24 with counters.
 - 2. Make the number 30 with counters.
 - 3. Which is more?
 - 4. How many more?
- B.
 - 1. Make the number 20 with counters.
 - 2. Which is more? 20 or 24?
 - 3. By how many?
- C.
 - 1. Which is closer to 24 - 30 or 20?
 - 2. How do you know?
- D.
 - 1. Draw a number line, and mark 24, 30 and 20 on it.
We say 24 is between 20 and 30.
- E.
 - 1. Is 30 quite near 24? Explain.
 - 2. Is 20 quite near 24? Explain.
 - 3. Which is a better approximation to 24?
 - 4. Give a number which is a better approximation of 24 than either 20 or 30.
 - 5. Use a number line to show why it is.

References

- Ashlock, R.B. Error Patterns in Computations: A Semi-programmed Approach. Columbus, Ohio: Charles E. Merrill, 1976.
- Bitter, G., & Mikesell, J.L. Activities Handbook for Teaching With the Hand-Held Calculator. Boston: Allyn and Bacon Inc., 1980.
- Bright, G.W. Estimation as Part of Learning to Measure. Chap. 5, Measurement in School Mathematics, 1976 Yearbook, National Council of Teachers of Mathematics, Reston, Virginia: NCTM, 1976.
- Carlow, C.D. Estimating (mimeo, 5 pp.) Toronto: Ontario Institute for Studies in Education, C. 1969.
- Carpenter, T.P., et al. Results and implications of the (1st) NAEP mathematics assessment: elementary school. Arithmetic Teacher, October 1975, 23(6), 438-450.
- Carpenter, T.P. et al. Notes from National Assessment: Estimation. Arithmetic Teacher, April 1976, 23(4), 296-302. pp. 206-302.
- Crawford, D.H. Difficulties and Errors in Computation--A Report and Commentary on some recent studies in England, Canada and the U.S.A." Paper presented at the 7th Annual Conference of the Research Council for Diagnostic and Prescriptive Mathematics, Vancouver, April 1980, (1980 a).
- Crawford, D.H. A Comparative Study of National and State (Provincial) Assessments of Basic Mathematical Skills in U.S.A., Canada and the United Kingdom. Paper presented at the Annual Conference of the Canadian Society for the Study of Education, Montreal, P.Q., June 4, 1980. (1980 b).
- Dienes, Z.P. Building Up Mathematics, (Rev. ed.). London: Hutchinson Educational, 1967.
- Fuys, D.J. & Tischler, R.W. Teaching Mathematics in the Elementary School. Boston: Little, Brown and Co., 1979.
- Guilbaud, G.T. Mathematics and Approximation. In Proceedings of the Third International Conference on Mathematics Education, Karlsruhe, 1976.
- National Council of Teachers of Mathematics. An Agenda for Action: Recommendations for School Mathematics of the 1980's. Reston, Virginia: NCTM, 1980.
- Payne, J.N. & Seber, R.C. Measurement and Approximation. Chap. 5, The Growth of Mathematical Ideas: Grades K-12, Yearbook of the National Council of Teachers of Mathematics. Washington, D.C.: NCTM, 1959.

Suydam, M.N. & Reys, R.E. (eds.). Developing Computational Skills. 1978 Yearbook. National Council of Teachers of Mathematics. Reston, Virginia: NCTM, 1978.

Trafton, P.R. Estimation and Mental Arithmetic: Important Components of Computation. Chap. 12, Developing Computational Skills, 1978 Yearbook, National Council of Teachers of Mathematics. Reston, Virginia: NCTM, 1978.

Underhill, R. Teaching Elementary School Mathematics. Columbus, Ohio: Merrill and Company, 1972.

Acknowledgment is hereby made of funding from the Faculty of Education, Queen's University, Kingston in connection with work on estimation and approximation.

Strategies for Diagnosing and Enhancing Learning of Place-Value Concepts

Bob Underhill, Virginia Polytechnic Institute
and State University College of Education, Blacksburg, Virginia

Place-value concepts have long been recognized as important elements in the development of understanding and the ability to apply and transfer number concepts. While several researchers have examined the merits of teaching base-and-place systems in experimental settings (Behr, 1975; Carpenter, McLeod, & Kkvarcius, 1975; Rathmell, 1972), only a few investigators have begun to study in detail the development of place value concepts among young learners and to examine inter-relationships, in the minds of young learners, among several mathematical capabilities.

Until recently, most exploratory work in mathematics was done in a clinical mode as by Piaget. Many researchers have also worked with individuals or small groups in pilot settings. For many years, Soviet researchers have used ascertaining and teaching experiments to study learning capabilities of youngsters. The purpose of an ascertaining experiment is to study what learners already know and to form conjectures concerning the types of experiences and the speed with which one might present new curricula. Once the data are available from the ascertaining experiment, one proceeds to the teaching experiment through which careful study of a number of subjects proceeds to determine the impact of formal instruction.

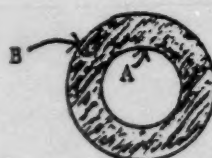


Figure 1. Zone of Proximal Development

In the teaching experiment, the researcher tries to determine how much learning can take place and often studies one or more concomitant variables which might influence this. Figure 1 presents the relationship. The interior of the inner circle, A, represents what the child already knows; this can be determined during the ascertaining phase. The interior of the outer circle, B, represents what is already known and what can be taught at this point in time; this obviously varies across subjects. Some conjectures are made about the shaded region during the ascertaining phase. The shaded portion represents "teachability" and is called the "zone of proximal development" by Soviet researchers (Vygotsky, 1962). The purpose of the teaching experiment is to gain insights about this zone of proximal development for a group of learners.

In the present work the investigator completed the ascertaining phase with a small group of first grade subjects (Underhill, 1972). Based on that early work and the work of Steffe (1976), a teaching experiment was conducted with a group

of youngsters to study relationships among place value concepts, conservation of number-related skills, and types of units counted by learners.

This paper, which presents information related to Phase I of a three phase project encompassing 1 1/2 years, is organized in three parts.

1. Selection of variables and instruments
2. Summary of experiences and results
3. Suggestions for teachers

Selection of Variables and Instruments

The investigator sought to identify independent variables which would have a bearing on children's development of early concepts and would make possible a careful investigation of the development of place value and regrouping concepts and their precursors. As a result of explorations through an ascertaining experiment, a review of research, and discussions with Steffe, two instruments were selected and two were created for Phase I:

Selected Instruments

Types of Units Counted (TEC), Independent variable
Quantitative comparisons (QC), Independent variable

Created Instruments

Place-value precursors (PVP), dependent variable
Base-and-place concepts (BP), dependent variable

Counting: Types of Units Counted

In Steffe's recent work with von Glasersfeld and Richards (1981), five stages in the development of counting were identified. Each stage is characterized by movement into a more abstract or removed-from-external-sensory-experience type of counting unit.

Children are said to count the following types of units: perceptual, figural, motoric, verbal, and abstract. The following two tasks will be used in describing counting behaviors of children who use these unit types.

Task 1. The subject is presented with the task displayed in Figure 2. There is a row of twelve chips; the first four are covered. The subject is asked to determine the total number of chips (twelve) when shown which one is the fifth from the left. (The four are not visible.)

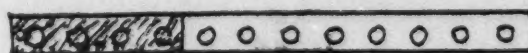


Figure 2. Task 1



Figure 3. Task 11

Task 11. The subject is presented with the task displayed in Figure 3. There is a card with two collections of chips, 4 and 3. The three chips are covered. The subject is told there are four and three, and is asked to determine the total number of chips, seven. (The three are not visible.)

In each of the counting unit types, the correct answer is not critical. The important criterion is the process or strategy which is used by the child who is counting. If counting is not used, then obviously no counting unit type inferences can be made based on that particular behavior sample.

Perceptual Type. The child who counts perceptual unit items always counts some type of external sensory data such as visual objects, sounds, or movements. The perceptual unit counter cannot answer Task 1 because there are no visible objects to count. Further, this counter when presented with Task 11 will (1) not respond, (2) count the visible items twice, or (3) use substitute sets for 3 and 4 (such as his/her fingers) but will treat the problem as one involving fingers rather than chips.

Figural Type. The child who counts figural unit items can use some imagination. This counter will imagine a set (often with the wrong numerosity, since s/he usually makes the judgment by spacing or dispersion). In Task 1, this counter will begin counting at the left (usually pointing or gazing at successive locations) and will proceed to count until s/he comes to the end of the covered portion and will then continue, counting the visible items. Example: (1, 2, 3,) 4, 5, 6, 7, 8, 9, 10, 11. In Task 11, this counter will imagine a set of three and will count imaginary locations. The counter will often respond correctly because of the availability of spatial patterns for small numbers (1-5).

Motoric Type. The child who counts motoric unit items has learned that imaginary items do not need locations; rather s/he has learned to depend on the motoric aspect of counting from perceptual and figural encounters. This counter may tap the table, tap a foot on the floor, count fingers under the table, or use some other motoric action to give each unit its "existence".

Verbal Type. The child who counts verbal unit items has further abstracted the counting process to where s/he believes numbers can be brought into existence simply by reciting the number names. In Task 1, this counter can simply say 1, 2, 3, 4, 5 and, relating five to fifth, 6, 7, 8, 9, 10, 11, 12. And, in Task 11, this counter can simply say (1, 2, 3) 4, 5, 6, 7.

Abstract Type. The child who counts abstract unit items is the most highly developed. In Task 1, this counter may relate fifth to five immediately and proceed "6, 7, 8, 9, 10, 11, 12." In Task 11, this counter may accept 3 as hidden and say "4, 5, 6, 7."

Determining the capabilities of individual counters is not as simple as it might appear since counters will generally gravitate to lower levels of counting unit types when they perceive that a task permits them to do so. The job of the investigator, then, is to be creative in the development of tasks to determine the highest level at which given subjects can succeed. This is often done through the use of problems with larger numbers (to decrease the likelihood of using spatial patterns) and story problems.

Quantitative Comparisons

Quantitative comparisons are made by children between two given sets which may differ by numerosity, density, and/or configuration. The Test of Quantitative Comparisons used in Steffe's University of Florida work (1976) was selected because the construct it measures was found to be highly correlated with mathematics achievement of first graders. The test selected is a shortened version of the one used in Steffe's earlier study (1970). The test consists of two warm-up items and eight other items. The reader is referred to Appendix A.1.1. of the 1976 source for a complete presentation of the instrument.

Place Value Precursors

A place value readiness test was developed by examining the grouping capabilities of several children of ages 4 through 6 (pre-school and kindergarten). Four basic types of test items emerged as a result of logical analysis and observation of the responses of young children:

- Subtest 1. Concrete objects using large groups (more than 30)
- Subtest 2. Verbal comparisons (no objects) using large groups (more than 30)
- Subtest 3. Concrete objects using small groups (four)
- Subtest 4. Verbal comparisons (no objects) using small groups (three)

Following is a detailed description of the Test of Place Value Precursors.

Warm-up. When the subject arrived, the Investigator used boxes and loose SOS pads, Ball jar lids, and birthday candles to help the subject use descriptive terminology such as "2 boxes and 1" or "1 box and 3". There was an initial tendency by all but one or two of the subjects to say "3" instead of "2 boxes and 1" and "4" instead of "1 box and 3", but after only slight coaching all but one of the subjects adjusted quickly to this manner of describing collections.

Subtest 1. A small box of raisins was opened and many raisins were poured out. Each item in this five item subtest consisted of two collections of boxes

and loose raisins glued to a large piece of tagboard. For each item, the subject was asked, "Are there more here or here or are they the same?" The order of "there", "here", and "same" were varied so as not to create a patterned response set.

Subtest 2. A 25 cent bag of M & M's was poured out and demonstrated to have many pieces of candy. The investigator placed two full bags and three M & M's on the table and said, "This is called two bags and three M & M's". Each item in this five item subtest consisted of an oral statement such as "Which have more M & M's or are they the same; 2 bags, or 1 bag and 3". Questions were repeated as often as needed and the researcher pointed out the difference between "5 M & M's" and "5 bags of M & M's" if a subject used the labels incorrectly.

Subtest 3. A miniature Halloween package containing only four Lifesavers was torn along the side and the investigator pointed out both visually and orally that there were four in the package. Each item in this six item subtest consisted of two collections of packages and loose lifesavers glued or taped to a large piece of tagboard. For each item, the subject was asked, "Are there more here or here or are they the same (or a variation of this statement)?"

Subtest 4. A miniature Halloween box containing only three Milk Duds was opened and the contents poured into the investigator's hand. It was pointed out both visually and orally that the box contained three Milk Duds. The investigator placed one full box and two loose Milk Duds on the table and said, "This is called 1 box and 2." Each item in this six item subtest consisted of orally presented questions such as, "Tell me in each question which is more or if they are the same ... 1 box or 2 Milk Duds?"

Base-and-Place Concepts

A fourth test was administered only to the experimental subjects. The purpose of this instrument was to determine the range of experimental concepts and skills mastered by each subject. This instrument consisted of eight items characterized as follows: addition with and without regrouping with and without objects, and subtraction with and without regrouping with and without objects. All items used groups of four and five. The specific items are presented later.

Summary of Experiences and Results

Other researchers have been unsuccessful in demonstrating the advantages of teaching non-decimal base-and-place systems. The obvious merit of teaching non-decimal base-and-place systems must be demonstrated to achieve greater gains or better understanding or greater transfer to "traditional tasks", e.g., computation. With these two pieces of information in mind, the researcher felt that the success of this study must be measured against three criteria:

1. Children can master the curriculum experiences,
2. Relationships which affect success can be identified, and
3. The research must produce evidence that the curriculum experiences are desirable.

Criterion 1 was addressed by demonstrating subject mastery of some or all of the planned experiences. Criterion 2 was addressed by studying subject performances as they relate to the two dependent variables. Criterion 3 was addressed by using a set of control subjects. The researcher believed that failure to address criterion three could easily result in the creation of a set of experiences which could be perceived by others as neither necessary or sufficient. The fact that children can master concepts or skills not in the curriculum is hardly an adequate rationale for including them!

Subjects

The treatment and control groups consisted of 12 first-grade children from one elementary school in a university community. The school is the most diverse in the community with a good cross-section of university, professional, business, and laborer families. In the fall of 1981, the researcher met at an open-house with many parents whose children were in the class. Subsequently, permission slips were sent home and twelve of the 21 sets of parents or guardians agreed to allow their children to participate. The 12 subjects were grouped into experimental and control groups based on preassessment data.

Distribution of counting unit types. Based on personal discussions with Steffe, the counters were partitioned into three categories:

1. Perceptual and Figural: 9 subjects
2. Motor and Verbal: 2 subjects
3. Abstract: 1 subject

Since counting unit type was deemed an important study variable, the researcher decided to group the subjects as follows:

Experimental subjects (X)

- Group 1: 3 subjects
- Group 2: 2 subjects
- Group 3: 1 subject

Control Subjects (C)

- Group 1: 6 subjects

The rationale for this grouping was that it provided groups of at least three subjects on which treatment effects could be studied. This seemed more desirable than the alternative of trying to make statements of relationship on the basis of one or two subjects. Thus, at the conclusion of the study, this arrangement would lend itself to some credible conjectures about treatment/non-treatment effects among Group 1 subjects and treatment effects between Group 1 and Group 2/3 subjects. Thus, the six experimental subjects consisted of all Group 2 and Group 3 subjects and three subjects selected randomly from Group 1 subjects.

Distribution on Test of Quantitative Comparisons

The performances on the eight-item test ranged from 1 to 8 among the twelve subjects:

Xs: 8 7 7 7 5 1

Mean = 5.8

Cs: 7 6 5 4 1 1

Mean = 4.0

In a check of the reliability, KR21 = .83 for the entire group. Table 1 presents the relationship between counting unit types and TQC scores. The Pearson product moment correlation coefficient between TQC and Counting Unit Types was only .45. The researcher decided not to use TQC scores as a stratification variable since this would have necessitated using at least one of the Group 2 or Group 3 counter unit type subjects as a control subject.

Table 1

<u>Relation between Counter Types and TQC Scores</u>								
	1	2	3	4	5	6	7	8
Group 1	XCC ^a			C	C	C	XXC	
Group 2					X		X	
Group 3								X

^aX = Experimental Ss, C = Control Ss

Schedule

The investigator met with the six Xs for 30 minutes on Monday, Wednesday, and Friday mornings at either 9:00-9:30 or 10:30-11:00 as school and personal schedules would permit. The lessons were designed to develop grouping concepts, first without the use of symbols then with them. Considerable emphasis was placed on counting. Toward the end, the special symbols were incorporated into addition and subtraction problems. During the two months of Phase I, the sequence of activities was as follows:

1. Two days of individual, videotaped preassessment
 - 6 Xs TUC, QC, PVP
 - 6 Cs TUC, QC, PVP

2. Four days of videotaped teaching
6 Xs
3. One day of individual, videotaped process assessment
6 Xs
4. Four days of videotaped teaching
6 Xs
5. One day of individual, videotaped process assessment
6 Xs
6. Four days of videotaped teaching
6 Xs
7. Two days of individual, videotaped post-assessment
6 Xs PVP, BPC
6 Cs PVP

Curriculum Experiences--Lessons 1-4

Objectives

- (1) Describe sets of objects in terms of groups and ones
- (2) Count the objects in a set in terms of groups and ones (Ex. 1 group of four and 3)

Day 1

1. Count macaroni onto a small paper plate, then dump the macaroni into a paper cup when a group of four is made. Say: 1, 2, 3, 4; 1 group of four. 1, 2, 3, 4; 1 group of four. Continue in the same way for up to 7 groups.

2. Keep track of groups. Same procedure as before, except count the groups of objects as they are formed. Say: 1, 2, 3, 4; 5; 1 group of five. 1, 2, 3, 4, 5; 2 groups of five. 1, 2, 3, 4, 5; 3 groups of five. Procedure continues up to three groups of five.

Day 2

1. Use transparent plastic cups and 1 cm plastic cubes. Show me ____ groups of ____ and ____.

2. Use packages of gum (5 sticks per pack). First count gum by packages (1 package, 2 packages, 3 packages, etc.). Then put out combination of packs and sticks and count (1 pack and 2 sticks, 2 packs and 3 sticks, etc.).

3. Count by packs and ones; trade 5 sticks for a pack.

4. Teacher builds sets of blocks using cups and loose blocks; learners describe.

5. One child builds a set; another describes.

Day 3

1. Count boxes of raisins and boxes and loose raisins. Ask "How many?" when presented with a collection.

2. Children use cubes, cups, and paper plates to build orally specified collections. Children make collections and describe.

Day 4

1. Use transparent plastic cups and metal washers. Count together to 2 groups of six. "Show me 1 group of six and 2."

2. Use washers in groups of six held together with a twister seal. Supplement and delete groups only. "How many?" (3 groups of six) Supplement and delete groups and ones. "How many?" (3 groups of six) Supplement and delete groups and ones. "How many?" (2 groups of six and 4) "Show me ____ groups of six and ____." "Let's count in groups of six."

Curriculum Experiences: Lessons 5-8

Objective: Use new symbols to describe groups and ones.

Day 5

1. Use washers and cups. Count in unison.
2. Use cups and clothespins. Count in unison.
3. Use fingers. Count in groups of 5 and groups of 4.
4. Introduce  for group of 3 and  for group of 4. Hold up a group of clothespins and match them to the appropriate shape. Then use , , , ,  to represent groups.
5. Use washers and cups to match numeral cards.

Day 6

1. Use styrofoam cups and 1-cm cubes. Review the work of the previous day.
2. Introduce  for groups of five and , , etc. Practice making groups of 3, 4, and 5 when shown numeral cards.
3. Introduce notation for groups and ones:  1,  3, etc. Practice reading numeral cards.
4. Use washers and twister seals in groups of six without written symbols. Supplement and delete the set; ask for descriptions (2 groups six and 5).

Day 7

1. Use numeral cards and clothespins. Build and describe collections. Read numerals.
2. Use a piece of tagboard for each child with large circles on the left side and a collection of blocks on the right side. Place blocks in circles when asked orally, "Make 2 groups of four."
3. Use tagboard and blocks. Extend to groups and ones: "Show me 2 groups of five and 3."
4. Repeat the previous activity with numeral cards.

Day 8

1. Provide each child with a set of numeral cards. "Show me the card which says 1 group of five and 4."
2. Teacher creates a set with cups and cubes. Learners choose a numeral card.
3. Practice unison counting. One child uses cups and cubes; all the children count together.
4. Practice counting with packs and sticks of gum.

5. Work in pairs. One child selects a numeral card. The other child constructs a set with blocks and cups.

Curriculum experiences: Lessons 9-12

Objectives: Develop semi-concrete skills, and develop addition skills

Day 9

1. Provide a worksheet. Match numerals ( 2) with dot patterns ( ..).
2. Provide a worksheet. Given dot patterns, fill in appropriate numerals ( _____).
3. Provide a worksheet. Given numerals, draw corresponding dot patterns.

Day 10

Two children had missed two lessons recently, and one other had a hard time on Day 9. The investigator met only with these three.

1. Use sets of washers and ones. Supplement and decrease; children give oral descriptions.
2. Use 1-cm cubes, some in baggies and some loose. Repeat previous activity.
3. Provide groups and ones drawn on paper. Children give oral descriptions.
4. Provide a worksheet. Given sets, children fill in numerals ( _____).
5. Provide a worksheet. Given sets, children write numerals and draw appropriate shapes.
6. Provide a worksheet. Given numerals, children draw sets.

Day 11

1. Present problems written on cards, one two-digit (2D) numeral and one one-digit (1D) numeral, no regrouping. Demonstrate with blocks and cups.
2. Use blocks and cups. Practice oral descriptions, supplementing and deleting.
3. Demonstrate $2D + 1D$, no regrouping, with cups and blocks. Provide a worksheet. Children find sums of $2D + 1D$, no regrouping, without objects.
4. Provide a worksheet. Children find some of $2D + 2D$, no regrouping, without objects.

Day 12

1. Use Cuisenaire Rods: light green, purple, yellow, white. Learners play with the rods and make length comparisons.
2. Children orally describe collections of rods (2 groups of five and 3).
3. Present two collections of rods. Children state sums without written problems.
4. Provide a worksheet. $2D + 2D$, rods, no regrouping, using 3 and 4.
5. Provide a worksheet. $2D + 2D$, rods, no regrouping, using 4 and 5.
6. Provide a worksheet. $2D + 2D$, no rods, no regrouping, using 3, 4, 5.
7. Provide a worksheet. $2D$ subtraction, rods, no regrouping, using 3, 4, 5 (only 3 people got to this task, the top counters).

The treatment was terminated at this time because the lower level counters were getting frustrated and seemed to be approaching their developmentally bounded limits. The last two lessons were laborious experiences for the low level counters. They found it hard to attend to the tasks and exhibited a substantial amount of avoidance behavior when asked to attend to the addition instruction.

Perceptual and Figural Unit Counters. The mean change scores for the six Cs and three Xs on the Test of Place Value Precursors are presented in Table II. The Cs exhibited some growth on Subtest 1 due to maturation and experience, but the overall mean change was only .5. The Cs exhibited growth on Subtests 1, 2, and 4 with an overall mean change of 4.0. The greatest changes between the Cs and Xs was evidenced on Subtests 2 and 4, the oral subtests.

Mean Changes between Pre- and Post-Assessments
on Test of Place-Value Precursors

	Subtest 1	Subtest 2	Subtest 3	Subtest 4	Total
3 Xs	+1.0	+1.0	+0.3	+1.7	+4.0
6 Cs	+0.7	-0.2	0.0	0.0	+0.5
Differences	+0.3	+1.2	+0.3	+1.7	+3.5

There was very little change within the Cs, and the changes within the Xs were generally distributed over the subtests. It seems fairly safe to conjecture that the changes were due to the treatment.

Treatment Differences on Group 1 vs. Groups 2 & 3 Counters.

The results of the eight item post-assessment on the experimental treatment are presented in Table 3. The criterion for a check in the table was that a learner could perform the task spontaneously or with a little interaction.

All but one of the Xs could add and subtract with and without regrouping in groups of four and five when using Cuisenaire Rods. In general, only the top counters could add and subtract with and without regrouping without manipulatives. Three of the six Xs could add and subtract with regrouping in groups of four and five abstractly; all three of them were Group 1 and Group 2 counter types.

Table 3
Item Performance of Treatment Ss by Counting Unit Group
on Base and Place Test

Items								
Subjects	1 ^a	2	3 ^a	4	5 ^a	6	7 ^a	8
<u>Groups 2 & 3</u>								
1.	x	x	x	x	x	x	x	x
2.	x	x	x	x	x	x	x	x
3.	x	x	x	x			x	x
<u>Group 1</u>								
4.	x		x		x		x	
5.	x	x	x		x	x	x	
6.			x		x			
	$\begin{array}{c} \boxed{1} \\ \boxed{2} \end{array} 2$	$\begin{array}{c} \boxed{2} \\ \boxed{1} \end{array} 1$	$\begin{array}{c} \boxed{4} \\ \boxed{1} \end{array} 3$	$\begin{array}{c} \boxed{5} \\ \boxed{1} \end{array} 3$	$\begin{array}{c} \boxed{1} \\ \boxed{2} \end{array} 3$	$\begin{array}{c} \boxed{1} \\ \boxed{1} \end{array} 1$	$\begin{array}{c} \boxed{4} \\ \boxed{1} \end{array} 1$	$\begin{array}{c} \boxed{3} \\ \boxed{1} \end{array} 1$
	+ $\begin{array}{c} \boxed{2} \\ \boxed{2} \end{array}$	+ $\begin{array}{c} \boxed{1} \\ \boxed{1} \end{array}$	- $\begin{array}{c} \boxed{1} \\ \boxed{2} \end{array}$	- $\begin{array}{c} \boxed{1} \\ \boxed{2} \end{array}$	+ $\begin{array}{c} \boxed{2} \\ \boxed{3} \end{array}$	+ $\begin{array}{c} \boxed{1} \\ \boxed{3} \end{array}$	- $\begin{array}{c} \boxed{1} \\ \boxed{3} \end{array}$	- $\begin{array}{c} \boxed{1} \\ \boxed{2} \end{array}$

^a These were done with Cuisenaire Rods. The other four were abstract.

Table 4 presents the final interesting Phase I finding. The subjects were rank ordered on the basis of the investigator's experiences with them and on a review of the dependent variable data. It appears that success is related to both of the independent variables.

Table 4

Rank Ordering of Performance

Rank	TQC ^a	Counter ^b
1	Hi	Hi
2	Hi	Med
3	Med	Med
4.5	Hi	Lo
4.5	Hi	Lo
6	Lo	Lo

^aHi = 7.8; Med. = 5; Low = 1

^bHi = abstract; Med = motor/verbal; Lo = perceptual/figural

Summary

Based on Place Value Precursors data presented in Table 4, it can rather safely be assumed that the treatment had a positive effect on the acquisition of the treatment skills and concepts. The following conclusions can be made rather definitely:

First graders can learn to--

1. count in groups and ones;
2. give numerical descriptions of groups and ones;
3. use a new symbol system for describing groups and ones;
4. subtract without regrouping using special symbols and manipulatives;
5. add without regrouping using special symbols and manipulatives;
6. add with regrouping using special symbols and manipulatives;
7. subtract with regrouping using special symbols and manipulatives.

What is quite remarkable in these results is that there was no instruction on addition or subtraction with regrouping in any of the twelve lessons! It appears that abstract problems cannot be solved by Group 1 counter types without specific instruction.

Suggestions for Teachers

The results of Phase I of this three phase study clearly indicate that counter-unit type and the conservation-of-number-related tasks (quantitative comparisons) are related to children's abilities to master place value-related concepts and skills.

Diagnosis

The task is to determine the counter-unit type, ability to make quantitative comparisons, and the facility of the learner with place value precursors. Once these data are available, one can plan an appropriate learning sequence.

Counter-Unit Type. The tasks and descriptions presented earlier in this paper can be used to determine in which of the three major groupings the learner lies. The important thing to remember is that learners will generally use the most primitive unit available to respond to a task. For example, the abstract counter will often count objects when they are available. Thus, as you are sure of a learner's ability to use a certain type of unit, ask questions which are more difficult, e.g., which cannot be answered with only this type.

Quantitative Comparisons. The Test is not needed to get a rough idea. Simply line up two rows of chips of equal length on a piece of tagboard, one row with six chips, one row with seven chips. Ask, "Which row has more chips or are they the same?" The non-conserver will say they are the same. On a second piece of tagboard, make two circles of different diameter, eight chips in each. The non-conserver will generally say that the larger circle has more chips.

Place Value Precursors. Using the guidelines for the Warm-up and Subtests 1, 2, and 4 described earlier, determine whether learners can think differently about collections and single items. For example, on a piece of tagboard glue 1 little box of raisins and four loose raisins on the left; on the right, glue two small boxes and one loose raisin. Learners who are weak on precursors will usually choose the raisin collection on the left.

For Older Learners. For children beyond the first half of the second grade, one additional task is particularly revealing of diagnostic information. This one can replace the precursors tasks. The purpose is to review flexibility of thought as it relates to place value. First, have the child count by ones as you place small blocks on the table. Second, help the child count by 10's as you place ten-rods on the table. Finally have the child count in increments and decrements of tens and ones as you add and take away tens or ones to a collection. For example, you might start with 34, then place a one (35), then another one (36); remove a ten (25), then remove another ten (15); place a one (16). If a child can perform this task, he/she usually has a pretty good grasp of place value.

Concept Enhancement

Group 1 counter-types who are also low quantitative comparers are probably not ready for place value instruction. They can begin to focus on experiences designed to teach groups and ones, but only a small amount of progress can be expected. Only experiences with manipulatives are appropriate, and should probably be limited to the range of concepts and skills covered during daily lessons 1-8 presented earlier.

Group 1 counter-types and transitional or high quantitative comparers can benefit from the use of manipulatives but cannot function well without them. The activities described in daily lessons 1-12 are appropriate.

Group 2 & 3 counter types who are transitional or high quantitative comparers (none in the study were low) can handle all the activities in daily lessons 1-12 with and without manipulatives.

Transfer

A final word--one clear outcome from this work is that the experiences do not transfer automatically to groups of 10. When the sequence of lessons is followed, it is recommended that they be followed with work in groups of ten, using 3 4 for three tens and four ones. After quite a bit of work with this symbol, then work needs to be done on the relation between 4 2 and 42.

When the learners are Group 2 or 3 counters, the set of experiences can be followed with regrouping in this sequence:

1. Regrouping for addition, e.g., $\boxed{2} \ 7 = \boxed{3} \ 2$
2. Addition with regrouping
3. Regrouping for subtraction, e.g., $\boxed{3} \ 1 = \boxed{2} \ 5$
4. Subtraction with regrouping

Footnote

(1) The investigator is deeply indebted to Mrs. Kay Olson from whose first grade class the teaching-experiment subjects were drawn. Her flexibility cannot be overstated.

BEST COPY AVAILABLE

References

- Behr, M.J. & Eastman, P.M. Interactions between structure-of-intellect factors and two methods of presenting concepts of modulus seven arithmetic--a follow-up and refinement study. Journal for Research in Mathematics Education, 1975, 6, 150-157. January, 1978.
- Carpenter, T.P., McLeod, D.B., & Skvarcius, R. A comparison of two levels of guidance and two levels of abstraction in presenting numeration concepts to prospective elementary teachers. Paper presented at the Annual Meeting of the National Council of Teachers of Mathematics, Denver, April 1975.
- Rathmell, E.C. The effects of multibase grouping and early or late introductions of base representations on the mastery learning of base and place value numeration in grade one. (Doctoral dissertation, University of Michigan, 1972). Dissertation Abstracts International, 1973, 33, 6071A. (University Microfilms No. 73-11, 237).
- Steffe, L.P. Differential performance of first-grade children when solving arithmetic addition problems. Journal for Research in Mathematics Education, 1970, 1, 144-161.
- Steffe, L.P. Quantitative comparisons and class inclusion as readiness variables for learning first grade arithmetic content. PMDC Technical report No. 9, 1976. (ERIC Document Reproduction Service No. ED 144 808).
- Steffe, von Glasersfeld, E., & Richards, J. Children's counting types: philosophy, theory, and case studies. (Draft copy). University of Georgia, 1981.
- Underhill, R.G. Teaching Elementary School Mathematics. Columbus, Ohio: Merrill, 1972.